Optimization Methods for the Smart Grid

Report commissioned by the

Conseil Français de l'Énergie

French Committee of

the *World Energy Council*

Michel De Lara, Pierre Carpentier, Jean-Philippe Chancelier, Vincent Leclère

École des Ponts and ENSTA ParisTech, France

This report includes invited contributions
from industrial and academic partners,
who are not engaged by its content

April 21, 2016

Abstract

The world's energy landscape is changing fast. Three key drivers are remolding power systems: renewable energies penetration, expansion of markets and of new players, deployment of telecommunication technology and smart meters. This report highlights how these changes impact and challenge the mathematical discipline of optimization. It displays interactions between the energy sector and the academic community, with testimonies from both parts.

More precisely, the report lays out optimization models and methods especially adapted to the smart grid paradigm. This paradigm is characterized by i) a growing number of decentralized and intermittent means of production (like solar and wind energies, the possible irregularity of which can complicate production management), ii) the deployment of *smart meters* and the development of automation of electric storage and electricity consuming devices (these intelligent and communicating equipments being supposed to make the electric consumption more controllable, hence more flexible). We develop to some extent different mathematical formulations of optimization problems, especially under uncertainty, because they open the way for different resolution methods. We dicuss several case studies.

We have chosen a progressive and pedagogical exposition, mixing point of views of the energy sector and of the academic community, with the goal to make collaborations easier, especially regarding stochastic optimization.

Contents

Scope of the Report

As Marcel Boiteux, honorary president of Électricité de France, puts it: *optimizing is obtaining the best compromise between needs and resources.* For electricity generation management, the problem is optimizing the use of existing resources (hydroelectricity, gas, renewables, contracts, etc.) to meet the electricity demand at the lowest socio-economic cost. Optimization and energy maintain a mutually beneficial relation since decades. As needs and resources change and complexify, research is triggered and new optimization methods find their place in the toolbox of engineers. Nowadays, the *smart grid* paradigm is starting to exert a strong pressure on the optimization discipline.

This report highlights the interactions between the changes in the energy sector and the mathematical discipline of optimization. It comprises six chapters (in English) and an appendix document (in French).

In Chapter 1, we highlight the changes in the world's energy landscape and how they start to impact the optimization of power systems. In Chapter 2, we relate the testimonies of energy actors that express explicit new demands to the optimization community. In Chapter 3, we underline the growing need to incorporate uncertainties in the framing of optimization problems, and the many ways it can be achieved. In Chapter 4, we lay out a panorama of classical and promising optimization resolution methods. Some ongoing works and practices are displayed in Chapter 5, with a focus on international experiences. We conclude in Chapter 6.

In a separate document in French, we yield an analysis of the scientific literature combining smart grids and optimization. This work is mainly composed of the internship report of Axel Ringeval, student of the *Économie du Développement Durable, de l'Environnement et de l'Énergie* Master (2012-2013). It benefited from the preliminary work of Fiona Foucault, student of the *Renewable Energy Science & Technology* Master (2011-2012). It has been completed with a few additional analysis by Adrien Cassegrain, student at École Polytechnique (2013). We thank the three students for their contributions, with a special mention for the substantial analysis of Axel Ringeval.

Chapter 1

Power Systems Undergo a Deep Remolding

The world's energy landscape is changing fast, and this will impact the optimization of power systems. In §1.1, we turn the spotlight on new issues that emerge, like renewable energies, storage, markets for reserves, demand side management, multiplicity of actors, etc. In §1.2, we point out three key drivers that are remolding power systems: renewable energies penetration, expansion of markets and of new players, deployment of telecommunication technology and smart meters. With the help of students, we have screened the scientific literature combining smart grids and optimization; we provide a synthesis in §1.3, whereas the full report is available as an appendix (in French). Finally, we try to delineate in §1.4 what the reader will and will *not* find in the report, in what concerns the mathematical problems and tools covered.

1.1 Snapshots on New Issues for the Energy Generators, Retailers and Regulators

Here, we simply turn the spotlight onto new kinds of problems for the energy generators, retailers and regulators.

Till now, energy production was rather controlled, predictible, secure, etc. because it mainly came from stocks (water reservoirs, fuels, nuclear fissile material) and the decision-maker could fix the quantities produced in a rather deterministic ways in order to meet a stochastic but rather predictible consumption. Uncertainty in production mostly came from failures of generation units. Nowadays, an increasing part of energy production comes from unpredictible sources, like renewables (wind, solar).

The load problem — how to anticipate and to respond to a variable demand — is complicated by the fatal and intermittent production of renewables (wind, solar), that can be seen as a negative consumption, changing the shape of the loads and making it more erratic. How to operate a generation mix with less base production and more fatal and intermittent resources?

With smart systems, the power system will display more and more capacity for flexibility, extending the options of all decision-makers. How to use this extension for more efficiency and for a better integration of renewables? What are the algorithms that have to equip smart controlers in order to reduce the risk of blackouts? Or to "shave the peaks" and optimize demand side management?

If the residents of a neighborhood return home by evening and all plug their electric vehicles to load their batteries, the electrical network will certainly "melt down", having not been designed for such configuration. How can we coordinate electric vehicles battery load? How can such coordination reduce the investments in network capacity expansion?

To buffer the intermittent production of renewables (wind, solar), we will certainly witness a growth in storage solutions (batteries, pumping station, etc.). What is the optimal management of the couple renewable-storage (aggregator or virtual power plant)? When is it better to sell renewable production directly to the market than to store it, expecting better profits later?

Electricity markets are expected to provide an equilibrium between supply and demand at least cost. Their rules will surely evolve, with more actors, of smaller size, and managing portfolios including renewables. Which market designs are adapted to this new context? What are the rules to fix the size of reserves, devoted to compensate variability in the production of renewables?

The traditional objectives of a power system — efficiency (in equalizing offer and electricity demand) and security — are also changing with new goals: enhancing energy efficiency, integrating renewable energies, reducing CO2 emissions and other pollutions, etc. How to achieve multiple goals, especially under uncertainty? What incentives provide to actors in a context of multiplication of decision centres and decentralized decisions?

1.2 Three Key Drivers are Remolding Power Systems

We highlight three key drivers that are remolding power systems: renewable energies penetration, expansion of markets and of new players, deployment of telecommunication technology and smart meters. The nature of those drivers boosts the demand for optimization in specific directions that we briefly touch upon, preparing a more elaborate discussion in Chapters 3 and 4.

1.2.1 Renewable energies penetration and environmental preoccupations

The European Union climate and energy package materializes the growing concern for the global impact of energy on the environment. It lays out three 20-20-20 objectives for 2020

- a 20% improvement in the EU's energy efficiency;
- a 20% reduction in EU greenhouse gas emissions from 1990 levels;

- raising the share of EU energy consumption produced from renewable resources to 20%.

All around the world, we contemplate the rise of renewable energies — at a large scale like onshore and offshore windpower, solarfarms, or more local like photovoltaics, heatpumps — and the importance of environmental constraints (CO2, nuclear risk, land use).

The consequences for optimization are manifold.

Renewable energies are intermittent and highly variable, and their availability is less predictable and less observable than traditional stock energies like fossile fuels or nuclear power. By substracting these fatal energy sources from the load curve, we obtain a new erratic load curve. This load curve has to be fulfilled by the remaining traditional energy sources (nuclear, fuels, water reservoirs) which now have to be able to respond to such variability. These features stress the need for stochastic optimization to handle uncertainty within the optimization process.

This unpredictable availability will trigger the development of energy storage at all ranges, especially small (batteries) and medium (pumping stations). Now, when you have storage, you have stocks, and when you have stocks, you have stocks dynamics and intertemporal trade-offs: when is it optimal to fill or to empty your stocks? This is why dynamical optimization will be more and more sollicitated.

Moreover, in presence of uncertainty, it is crucial to distinguish those decisions that make use of available information from those that do not. The prototypical situation is the so-called two-stage recourse problem: at the first stage, decisions are made knowing only a statistical description of the randomness, whereas the second decisions are based on the observed values of the uncertainties. The structure of so-called day ahead markets reflects this prototypical situation: units have to be mobilized day $D-1$ — on the sole basis of statistical knowledge of uncertainties of day D like failures or demand (itself strongly driven by weather conditions) — whereas productions levels are adjusted and reserves are possibly called on day D in function of uncertainties, then revealed and known. This is why stochastic optimization will be more and more sollicitated.

An important feature of renewable energies is their spatial dispersion: think of wind, sun, small hydropower. It is to be feared that not only stochastic dynamical optimization will be needed, but also decentralized optimization where local information is not systematically available to a central planner. This question is important for the design of electricity markets in the presence of many agents with very different capacities, like small and large producers, with different reaction times, different needs and evaluation criteria.

Environmental preoccupations go beyond the development of renewable energies. We give two examples.

The concern for the impact of energy on the environment can lead to technological breaks, allowing for industrial changes like the substantial development of electric vehicles. The impacts on the load curve can be dramatic.

Technical constraints on the power systems are increasing and complexifying due to society demands concerning pollution, health and ecological impacts, multiple uses of resources (land, water), safety. These constraints will translate into specific mathematical constraints in optimization problems, making them harder to solve like in the case of so-called chance

constraints.

1.2.2 The expansion of markets and of new players

The deregulation of the energy sector is followed by an expansion of markets and by the emergence of new players.

Nowadays, we witness the growing role of territories and cities as actors in the energy landscape, not to speak of individual customers (now coined "prosumers") being both clients and generators. These actors have an impact both on the demand (load) by means of energy efficiency programs, for instance, and on the generation, by fostering the development of local renewable energies and local storage (batteries, pumping stations).

Therefore, the structure of the local mesh in electricity networks is expected to become more complex. On the other side of the spectrum, the size of electricity networks is also changing. As an illustration, the Pan European system counts 10,000 buses, 15,000 power lines, 2,500 transformers, 3,000 generators, 5,000 loads. The supply equals demand equilibrium can now be performed by massive transfers from distant network nodes (like the storage of wind energy, produced in the South of Europe, in reservoirs to the North).

As a consequence, the whole decision structure is changing: production is more dispersed, less centralized but with pan-continental coordination centers; centers of decision are also being scattered, with multiple interacting players. Such a dispersion makes life hard for those in charge of ensuring that supply equals demand over their perimeters, be they the TSO, the transport network manager, or else.

We are yet contemplating an expansion of energy markets that need to integrate markets for a rising number of time scales (years, months, weeks, days, intra-days, minutes) and a rising number of agents (small and large producers and consumers).

1.2.3 The deployment of telecommunication technology and smart meters

Some say that we are witnessing the marriage between energy and telecommunication, materialized through the notion of smart grid (to be discussed more intensively and detailed in §1.3.2).

Indeed, infrastructures and technologies are changing and more and more special devices are installed in the electric grid: smart meters, new electric appliances, grid communication devices, information systems, storage (thermal, accumulators, etc.), electric vehicles, intermittent energies, etc.

In a nutshell, the notion of smart grid covers an infrastructure project (equipment hardware) together with promises to be fulfilled by a "smart power system" (optimization software).

Indeed, since the energy supply is less flexible — due to the dispersed and intermittent nature of renewable energies — the focus is put on making the demand more flexible. This is what lies behind the vocables of smart management, smart operation, smart meter manage-

ment, smart distributed generation, load management, advanced distribution management systems, active demand management, residential curtailment, distribution management systems, storage management, smart home, demand side management... Technology is expected not only to measure more finely (sensors) but to provide decision tools (controllers) to be able to influence the demand, which was not the case till now to such an extent.

Acting on the demand is certainly a major control lever, although it can be restrained by the consumers price-elasticity. Note that the demand becomes more and more erratic because it relies more on electricity, with electric heating and new uses, like bigger and bigger TV screens and electric vehicles.

In what concerns optimization, we retain from the smart grid paradigm the idea that it provides more observation (sensors) and control (controllers) possibilities that can be exploited to provide an improved allocation between needs and resources in a context of spatial dispersion and temporal variability of both energy sources and demand.

1.3 Analysis of the Scientific Literature Combining Smart Grids and Optimization

With the transformations in power systems, their management will become more complex. Screening the smart grids literature, the answer seems to lie in more "smartness". Not only does the grid have to be equipped with smart devices, it also needs smart control and optimization.

We yield now a summary of an analysis of the scientific literature combining smart grids and optimization to be found in a separate document in French. This work is mainly the internship report of Axel Ringeval, student of the *Économie du Développement Durable, de l'Environnement et de l'Énergie* Master (2012-2013). It benefited from the preliminary work of Fiona Foucault, student of the *Renewable Energy Science & Technology* Master (2011-2012). It has been completed with a few additional analysis by Adrien Cassegrain, student at École Polytechnique (2013).

1.3.1 Study scope

The study of the scientific literature combining smart grids and optimization tries hard to answer the following questions. What are the needs in optimization put forward in the related academic literature in smart grid? What are the types of problems tackled by the scientific community? How are these optimization problems formulated? What are the algorithmic and numerical resolution methods? The scientific publications that have been analyzed are, closely or remotely, in connection with the concept of smart grid.[1] However, only publications presenting problems of optimization have been retained.

[1] At the time of the students studies, we did not ask them to screen the scientific literature treating renewable energies management with optimization methods. We discovered later that this literature was substantial, with some interesting contributions.

The scope of bibliographical search included problems of management of production and of electric consumption, but excluded problems of electricity transport, like the plentiful literature on "optimal power flow".[2] Only optimization problems related to the production and supply of electricity have been considered in this study.

1.3.2 What is considered as smart grid in the literature review

The notion of smart grid is found in various disciplines, like electrical engineering, IT engineering, telecommunications, etc.. Many academic or industrial actors refer to this term nowadays, with different acceptations. The scientific literature helped us to bring to the foreground a focused notion of smart grid, especially thanks to two publications, remarkable for their didactic contents, that is, [36] and [1].

Intelligence in Electricity Networks for Embedding Renewables and Distributed Generation [36]

The article [36] does not deal directly with the notion of smart grid, but is an opportunity to define fundamental associated concepts, frequently encountered in the scientific literature: *distributed generation* (DG), *demand response* (DR), *distributed energy resources* (DER) and *virtual power plant* (VPP).

This publication begins with a clear description of the current functioning of the electricity industry and the evolutions ahead. The authors first outline the chain of value from production to consumption, both from physics and economic perspectives. Then, they underline that electricity has, for long, been produced by few centralized units with large unitary power. As renewable energies contribute more and more, they challenge the energy sector because they are widespread and intermittent, hence less controllable, with low unitary power. Resorting to more flexibility in the demand might be necessary to compensate for the lack of flexibility in production. However, this requires a capacity in coordinating individual consumers (to which the diffusion of smart communicating devices could contribute). Finally, the following fundamental concepts are defined:

- *distributed generation* denotes the production of electricity by units connected to the distribution network, or even to consumers directly;
- *demand response* refers to the possibility of modifying the consumption of installations or devices, especially at short time scales (almost immediately) in response to proper signals like prices;
- *distributed energy resources* refers to the triplet "distributed generation", "demand response" and *energy storage*;

[2] This type of OPF problem consists in minimizing the cost of transport of the electric production towards the places of consumption; it aims at determining the optimal characteristics of the tension — amplitude and phase shift — as well as the optimal powers — active and reactive — in every node of the network.

- a *virtual power plant* (VPP) is a portfolio of distributed energy resources controled by a central unit; a VPP can play the role of flexibility reserve and offer balance service to a network manager; VPP and *aggregator* are related notions.

Adaptive Stochastic Control for the Smart Grid [1]

The article [1] puts forward an insightful notion of smart grid. Indeed, by this term, the authors refer to a vast integrated system, the multiple Technological components of which they review, be they emergent or expected. They would considerably affect the current modes of energy production, transport, distribution and consumption. Concerning modes of production, they envisage for example a massive development of decentralized means of production (*distributed generation*), wind and solar production capacities — thermal as photovoltaic — coupled with important capacities of storage. Concerning modes of transportation and distribution, they imagine a network equipped with technologies allowing a more flexible electricity routing, and rather organized in the form of *microgrids*, with the potential to be insulated as they would benefit from decentralized means of production.

Finally, concerning modes of consumption, they defend the widespread idea that an advanced deployment of smart and communicating devices (*smart meters*) could create a vast potential of reduction and/or transfer of consumptions (*demand response*), what is more in a context of important penetration of battery-driven vehicles.

This almost utopian vision of a smart future network has the merit to make the reader sensitive to the multiple relative needs for optimization it requires. Under the vocable of *Adaptive Stochastic Control*, the authors refer to the ideal dynamic control that it would be necessary to develop in order to simultaneously manage operational problems, as currently is *optimal power flow*, but also problems of choice of investment, of maintenance planning, of optimal answers to emergency situations and others, conceivable or not yet imagined. All these optimization problems will become stochastic and will be made more complex by the introduction of real time electricity pricing mechanisms, by the possibility of delaying consumptions, by the multiplication of refillable battery-driven vehicles, storage capacities, intermittent production means, etc.

Then, the authors undertake an almost theoretical presentation of dynamic and stochastic systems. They define for example the notions of state, control, information variables, of function describing the dynamics or the objective of an optimization problem, etc. They also define the notion of strategy (*policy*) to introduce then what constitutes the scientific core of this article, namely the use of methods to select strategies based on various approximations of a value function by projecting it on a basis of functions. These methods are called "Approximate Dynamic Programming" and would allow to decompose large-scale problems — as would exactly be the ones relative to the management of smart grids — in order to solve them approximatively. However, they only apply their approach to a simple optimal power flow problem, where the question lies in determining the electric flows on a small network constituted by a source of stochastic consumption, of a free but stochastic electric production (solar or wind), of a connection to the network which constitutes an additional costly source with stochastic price, and of a storage capacity as a battery for example.

What is considered as smart grid in the literature review

In the light of these two publications, the smart grid paradigm is characterized by:

- a growing number of decentralized means of production, the possible irregularity of which can complicate production management;
- the deployment of *smart meters* and the development of automation of electric storage and electricity consuming devices; these intelligent and communicating equipments are supposed to make the electric consumption more controllable, hence more flexible.

New possibilities of management are then offered:

- the possibility of controlling the production on more local perimeters, as for example a portion of distribution network potentially insulated;
- the possibility of aggregating consumptions and, to some extent, of being able to delay them to reduce the consumers' invoices, to prevent consumption peaks, or to supply regulation services to a network manager.

The profit opportunities that will doubtless accompany these new activities could favor the emergence of new actors, like the various types of aggregators found in the literature.

1.3.3 Analysis table of the literature and main conclusions

Some basic mathematical questions have allowed to design an effective reading "filter" during the thorough work of scientific literature review. Many of these questions are a straightforward consequence of the pressures exerted by the remolding of power systems on research in optimization, as touched upon in §1.2 and as will be developed in §2.1.

- Is the optimization problem stochastic or deterministic? That is, are data taken into account as having single values known in advance (like the load of next day) or are they handled by means of distributions (temporal profiles of wind power, of load, etc. given by stochastic processes or by scenario trees)? As discussed in §1.2 and as will be developed in §2.1, the rise of renewable energies is a strong incentive to mobilize stochastic optimization.
- Is the optimization problem formulated in a dynamic or in a static way? That is, is the stress put on a specific dependence with respect to the time variable (taken generally at discrete steps or stages) as when stocks and reserves are under scrutiny? As discussed in §1.2 and as will be developed in §2.1, the rise of energy storage should foster dynamic optimization.
- Is a particular information structure taken into account in the optimization problem? As stressed in §1.2 and as will be developed in §2.1, production and decision centers are more and more spatially dispersed. This is why the nature of the information available to each decision center is a relevant input to frame optimization problems.

- Is the optimization problem a linear problem or not? Linear programming is a well-known and proved method to solve a large class of problems. More generally, convex optimization problems are amenable to powerful numerical algorithms. This is why such a screening is especially relevant when dealing with large-scale systems.

- Does the optimization problem include discrete variables, hence resorts (at least partly) to combinatorial optimization methods?

- Is the optimization problem solved by means of a specific method, or even a specific numerical algorithm? In many papers, the emphasis is more on the numerical method used than on the mathematical framing of the optimization problem. By resolution methods, we mean such various ones as linear programming, decomposition by dualization of different or specific constraints (time, space, scenarios), stochastic programming, dynamic programming, progressive hedging. We try hard in Chapter §4 to provide the reader with a synthetic panorama of resolution methods for multi-stage optimization problems under uncertainty.

We observed that the optimization problems identified in the scientific literature are mainly a matter of linear programming and, more generally, of convex optimization. They are concerned with various manners to model and to solve large-scale problems. Linear programming is the most common approach but, outside the linear case, some articles propose decomposition methods. In the report in French, we displayed them according to three categories: temporal decomposition, spatial decomposition or scenario decomposition.

1.4 What You Will and Will *Not* Find in the Report

First, we glimpse at how some authors tackle issues at the intersection of smart grids, optimization and stochasticity. Second, we sketch by contrast what will be our angle in the report.

1.4.1 A glimpse at issues often related to smart grids, optimization and stochasticity

Since we started in 2011 our review of the scientific literature combining smart grids and optimization (see §1.3), we observe a growing segment of literature explicitely tackling intersections between smart grids, optimization and stochasticity.

Smart grids, optimization

For instance, the journal *IEEE Transactions on Smart Grid* devotes a *Special section on optimization methods and algorithms applied to smart grid* in Volume 4, Issue 4, Dec. 2013.

Smart grids, stochasticity

For instance, Hao Liang and Weihua Zhuang claim, in *Stochastic Information Management in Microgrid Operations* (IEEE Smart Grid, April 2014):

> Given the randomness in renewable power generation, and the buffering effect of energy storage devices and various customer behavior patterns, stochastic information management — employment of stochastic modeling and optimization techniques for information processing and decision making — is critical for microgrid operations to ensure system efficiency, reliability and sound economics.

For the analysis, we agree with these authors. However, we do not always coincide with them and other authors in what concerns *tools*.

To be more specific, we focus on an interesting paper to be published in *IEEE Communications Surveys & Tutorials* (2014). In *Stochastic Information Management in Smart Grid*, Hao Liang, Amit Kumar Tamang, Weihua Zhuang and Xuemin (Sherman) Shen mention basic theories and techniques in stochastic modeling, optimization, and control:

- Convolution technique;
- Interval based technique (which includes Robust optimization);
- Moment estimation;
- Dynamic programming (which includes Partially observable Markov decision process, Approximate dynamic programming, Markov decision processes, Stochastic dynamic programming, etc.);
- Stochastic control (which includes Kalman-Bucy filtering, Model predictive control, etc.);
- Stochastic game;
- State estimation;
- Queueing theory;
- Stochastic inventory theory;
- Monte Carlo simulation.

We do not feel quite confortable with such lists that mix theories and techniques in disparate fields. Concerning tools, we will cover a smaller ground.

Optimization, stochasticity

In this report, we stress that multistage optimization under uncertainty is characterized by how two different issues are handled:

- risk, or how to perform a trade-off between all possible uncertainties;
- online information, or how to take advantage of the data (on the uncertainties) that is available at every stage before making a decision.

Regarding the techniques often pointed as relevant for dealing jointly with optimization and stochasticity, we have the impression that many of them only treat one of the above issues.

Whereas *sensitivity analysis* is a widely used technique to assess the impact of uncertainty, we do not discuss it. As far as we understand sensitivity analysis, it aims at comparing open-loop solutions. An "open-loop" solution to a multistage optimization under uncertainty indeed performs a trade-off between all possible uncertainties, but does it once and for all times (such as in a one-shot investment problem). By contrast, a "closed-loop" solution takes advantage of the information (on the uncertainties) that is available at every stage before making a decision. We refer the reader to [35] which makes an excellent job at enlightning modeling issues with stochastic programming, and specifically discusses the pros and cons of sensitivity analysis.

Although *Monte-Carlo* is often mentioned when optimization is performed under stochasticity, we find it hard to pinpoint a specific optimization method. By contrast, the *stochastic gradient* method is a well-established stochastic optimization method which makes use of Monte Carlo drawings inside a gradient method. It is relatively easy to implement, but it aims at providing "open-loop" solutions, and is thus not adapted to compute decision strategies.

As far as we understand *model predictive control* (MPC), this technique mixes system identification with rolling horizon deterministic optimization. At every stage, the current data is used to formulate a finite horizon deterministic optimization problem, and only the current stage optimal decision is implemented. Thus, MPC computes closed-loop solutions, but we are not sure that a trade-off between all possible uncertainties is performed.

Optimization, complexity

Most "real life" optimization problems are too complex to be numerically solved directly. We briefly list some of the many ways found in the academic literature to tackle complex optimization problems, pointing to well-known references, without aiming at exhaustivity.

Heuristic. We can look for heuristic solution, either by looking for the solutions in a more limited class of solutions — "open-loop-feedback-optimization" (see [12]), approximate dynamic programming (see [13, 51]), machine learning (see [31]), linear decision rules (see [37]) or more generaly parameterization of decision rules — or by cunningly trying to find a good solution through method like simulated annealing (see [59]), or genetic algorithms (see [29]).

Specific problems. We can also make some approximation of the problem itself, and make the most of some mathematical properties of the (approximated problem). For example, one finds very efficient algorithms for linear programming problems (see [21]), quadratic programming, semi-definite programming, conic programming, (see [81, 9, 2]) large classes of mixed integer linear programming (see [41]), etc.

Decomposition. Decomposition approaches (see [6, 56, 19]) consist in partitioning the original optimization problem into several *subproblems* usually coordinated by a *master problem*. We then solve each subproblem independently, and send the relevant part of the solutions to the master problem. The master problem then adjusts the subproblems, that are to be solved again, and so on. The numerical gain is contained in the fact that, if the original problem is of size S, solving N problems of size S/N, even with iterations, might be much faster than solving the original problem.

1.4.2 What you will find

This report is about *optimization*. You manage production facilities, you have to clear markets for energy, you wonder in which technologies you should invest, you want to reduce energy consumption or energy bills by shifting demand, etc. In all these situations and in many others, you have some room for manoeuvre, delineated by constraints, and you look for the "best" course of action or, at least, for a "good enough" one.

When we speak of *control*, or of *control theory*, we focus on methods to find our way among constraints. And there are generally many ways to satisfy constraints and to achieve goals materialized by constraints. Here, a course of action is "good enough" as soon as it meets the constraints.

But, when we resort to *optimization*, we lay ahead a special goal: optimizing a numerical criterion among all possible courses of action delineated by constraints, hence looking after the "best" course of action. To be more specific, this report touches *team theory* — where different players with different informations jointly contribute to optimize the same criterion — but not *game theory*, where each player optimizes his/her own criterion.

This report especially tackles *stochastic* and *multi-stage* or *dynamic optimization* problems. This is why Chapter 3 is devoted to how to incorporate uncertainties in the framing of optimization problems, and the many ways it can be achieved. The result of this crafting is an *optimization model*. Then, you can resort to mathematical *methods* to characterize solutions, and to *algorithms* to solve the problem numerically. This is the object of Chapter 4, where we lay out a panorama of classical and promising methods adapted to *multi-stage stochastic optimization problems*. To complete the picture, we invite international colleagues to share experiences, works and practices in Chapter 5.

Chapter 2

Energy Actors Express Renewed Demands towards Optimization

In Chapter 1, we highlighted the changes in the world's energy landscape and how they started to impact the optimization of power systems. Now, we progress towards a more specific description. In §2.1, we elaborate on the challenges for optimization theory and practice; with this, we pave the way for Chapters 3, and 4. In the following §, we relate new demands towards optimization that arise from the energy world. Indeed, we observe that a few energy actors express explicit new demands to the optimization community — like the US Department of Energy (DOE) and French small companies — especially in stochastic optimization, in a context of strengthening of innovation capacities.

2.1 Challenges for Optimization

In §1.1 and §1.2, we laid out examples of new problems that follow the remolding of power systems. In some occasions, we took the opportunity to stress the incidence for optimization. Here, we elaborate more systematically on the challenges for optimization theory and practice raised by this remolding.

As said in the introduction, optimizing is obtaining the best compromise between needs and resources. Compromise, needs and resources will be our guiding light in exploring what challenges lie ahead for optimization.

We will focus on electricity generation management, as opposed to "optimal power flow" types of problems (see footnote 2). The central problem is optimizing the use of existing resources (hydroelectricity, gas, renewables, contracts) to meet the electricity demand at the lowest socio-economic cost, at all times and at any node of a network.

As we have seen, the last few years have witnessed a tremendous increase in the installed capacity of renewable energy generators (wind, solar, etc.). Due to the variable and unpredictable nature of renewable generators, resources availability has become intermittent. Needs also are highly variable. Thus, *optimization is challenged by uncertainty*, more variable and at shorter time scales. How does the balance between needs and resources move

when input data are uncertain?

Needs are more complex. Indeed, beyond supplying energy, security is a major concern (breakout), as well as environmental preservation (CO2 emissions, pollutants, biodiversity). This will compel to introduce certain types of constraints (robust or in probability) and risk (risk measures) in the optimization framework. in extending constraints from a deterministic to a stochastic setting, many mathematical formulations are possible, leading to different problems. This will require to develop specific mathematical approaches and algorithms, depending on the formulation. Thus, *optimization is challenged by the multiplicity of objectives and how they combine with uncertainty.* How are the needs mathematically formulated when they bear upon uncertain outputs, corresponding to security and ecological requirements?

Concerning the best compromise between needs and resources, this may be the vision of a central planner, but it is now blurred by the multiplicity of actors having private information and goals. *Optimization is challenged by the multiplicity of actors* and by *decentralized information.*

Thus, in addition to the already difficult issues of *dynamics* (stocks variations) and *large scale* in the deterministic optimization of power systems, the task is now made more complicated by *uncertainty*, *multiplicity of objectives*, and *decentralized information.*

In deterministic optimization, linear programming allows to tackle rather large problems. Powerful methods and software allow to solve linear optimization problems with a large number of variables related by linear equalities or inequalities. Of course, the price to pay is a linear formulation of problems, may they be linear or not (hence resorting to linearization). Therefore, a natural avenue consists in incorporating uncertainty, multiplicity of objectives and decentralized information within a linear programming framework.

2.2 SETEC Energy Solutions (Nicolas Lebert)

Setec is an independent engineering company, which has competences in many fields such as infrastructure and transport systems, geotechnical and environmental management, water, energy and others. Setec Energy Solutions is a subsidiary created in 2011 in order to develop energy production and management projects. In all those projects, Setec employees can have a role of general consultant or project managers, provide basic and detailed engineering, and lead the feasibility studies or the due diligence. With this activity, Setec has an overall vision of the electricity field and its future evolution. Setec ambition is thus to stay on the edge of all the future evolution of the electricity market, and most especially of the integration of the renewable energies in the power grid.

One of the main aspects of a feasibility study is the economic part, especially the estimation of the return on investment of a project. When dealing with hydropower schemes, there are uncertainties involved in the revenue generated by the plant. In the case of run off river schemes, energy production optimization stays a simple problem, but when it comes to a chain of reservoirs, the decisions are much more complex: when should the plants produce electricity? What should be the levels of water? When should the energy be stocked? The hydropower schemes should moreover respect environmental constraints that may restrict

the possibilities. All those decisions have finally to be optimized in an environment where the water inflow is difficult to forecast at long term and the electricity price is very volatile.

In order to deal with those issues, Setec initiated cooperation with the Optimization and Systems Group at Cermics (École des Ponts ParisTech), which is still ongoing. Through this collaboration, Setec built a hydropower dedicated tool, *Hydroptim*, that optimizes and simulates chains of reservoirs and plants. The software optimizes the decision matrix of hydropower schemes in order to maximize the revenue depending on the water inflows and the prices. In a deterministic environment, with known water inflow and electricity price, Hydroptim is able to find an optimal strategy for small chains, and to approximate it with decomposition-coordination methods for big ones. In order to get more realistic results, a stochastic dynamic programming algorithm, which takes into account the water inflow uncertainty, was implemented with the help of the Cermics. The optimization of large chains of dams with this algorithm is although still limited because of the computational time it would require.

Another topic that will have to be considered in the future is the valuation of the electricity. With Hydroptim the valuation of the electricity is time dependent but not stochastic: this is enough for small hydro plants that beneficiate from fixed buying price, but this does not allow modeling a plant whose electricity is sold on the EPEX trading market. When the fixed price will not be possible anymore for small hydro plants, the uncertain electricity price will be a critical factor in the economic evaluation of a project. To include the price uncertainty in the model, there will be a need to develop a market price model, and to improve the optimization algorithm, in order to take it into account without increasing the run time.

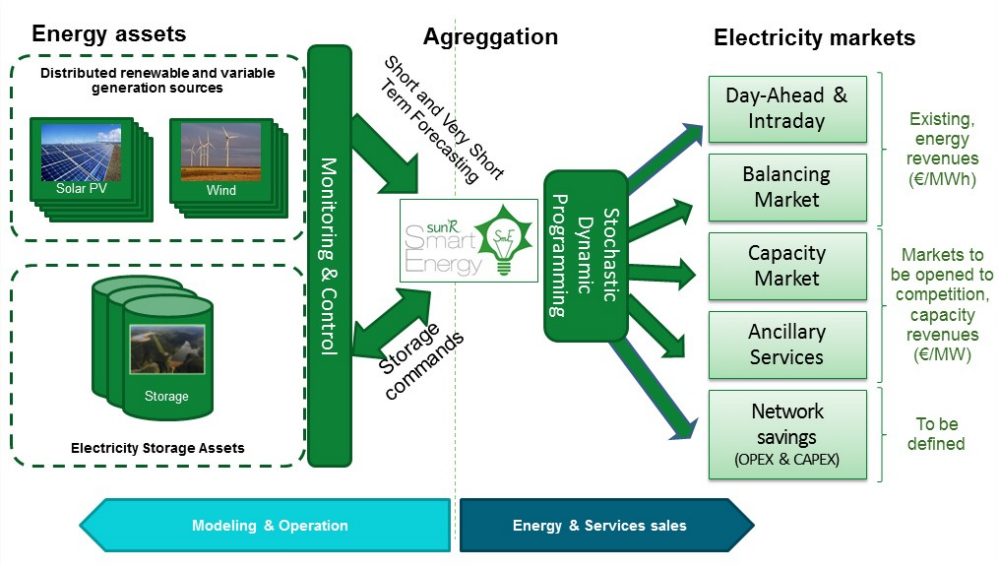
As the part of the renewables energies in the power mix is growing, the stability of the power system may be threatened. Setec is currently working on electricity storage solutions, such as pumped storage facilities, which will strengthen the power grid. The renewable energies will also strongly modify the EPEX prices, and may increase the occurrences of extreme values such as negative price. To predict the economic viability of pumped storage facilities, there is thus a need to predict the demand curve on the network as well as the market price, and to link them through an optimization tool.

In its activity, Setec regularly needs to update its vision of the power system, prices and electrical demand to optimize the economical profitability of its projects. The current trend is that the revenue estimation of assets will become more complex and will require powerful optimization tools to provide the best level of advice to decision makers.

2.3 SUN'R Smart Energy (Davy Marchand-Maillet)

2.3.1 Case study: Sun'R Smart Energy, RES+storage agreggator

In the SunHydrO project framework, Sun'R Smart Energy and its seven industrial and academic partners are aiming to facilitate the integration of variable Renewable Energy Sources (RES) into the electrical system by combining those with appropriate electricity



storage assets and particularly small and flexible pumped hydro storage (PHS) units.

Hence, the main project goal is to create a Virtual Power Plant (VPP) with

- solar PV power plants: some being built on farms roofs, other ground based, mounted on trackers and/or with a concentration technology (CPV),
- wind mills and
- flexible small PHS.

This VPP is designed to address the main issues with variable renewables, namely

- to have the RES production to participate to the electricity market, so RES producers can leave the "Feed-in Tariff" mechanisms,
- to smooth the variations of part of the RES generation to contribute to system stability, notably by contributing to ancillary services,
- to ease the connection of new generation facilities on the network without investing too much in wires or transformers.

To make sure that the contribution of the VPP is positive, the so-called aggregator faces multiple stochastic optimization problems as follows.

1. Strategic optimization
    - Optimization of the RES portfolio for a given PHS unit.
    - Conversely, dimensioning of an additional PHS unit in the VPP given the existing VPP and its supposed evolution.

2. Operational optimization

    - Determine the bids to be made on the markets given the stochasticity of the electricity prices and RES generation forecasts.
    - Optimize the running plans of the PHS in pump/turbine mode down to a 50kW accuracy, to fulfill the commitments of the aggregator.

3. Real-time optimization

    - Given the dispatch of the PHS unit, generate the actual operations, to comply with the specificities of the PHS (forbidden zones, optimize efficiency, etc.).

It is well known that RES variability is strongly influenced by the weather conditions. One of the specificity of the project is that the RES generation sites are equipped with meteorological meters that will allow the aggregator to benefit from premium forecasts of its RES output and its correlation with the national RES output and so prices on the markets. With this privilegied access to RES generation data, Sun'R Smart Energy will benefit from a significant competitive advantage versus competitors on the energy markets.

This integrated approach strongly requires optimizations methods and tools, because strategy, sales and operations are deeply connected and the proper actions require an optimized coordination.

Finally, the SunhydrO project is an initiative towards a more distributed energy system, because Sun'R Smart Energy believes that such systems are much more resilient than centralized ones. Hence, black-outs may become brown-outs (localized failures, with no impact on the European system if a failure occurs in Brittany) with a decreasing probability of occurrence.

2.3.2 Case study: agrivoltaics, solar panels positive to crops

Through the Sun'Agri research program, Sun'R contributes to the development of solar panels which are specifically built and equipped:

1. the panels are over-elevated so that farmers can grow plants beneath, for example grapes or vegetables,

2. the panels are mounted on 1-axis trackers, so that a given solar panel can be perpendicular to the sun rays and create some shades for the plants OR be aligned with the sun rays and let the light reach the plants; the angle can be intermediary, so that part of the light goes to the plants and the other part serves for solar PV generation.

This innovative approach, the "agrivoltaic concept", does not only solve the conflict of dedicating land to produce either food or energy. Hence, with an appropriate control algorithm, it has been proved that such a system will not only maintain the productivity of the crop but even enhance it, either on a quantitative basis or a qualitative one, while adding significant solar PV generation onsite and, last but not least, saving water.

Thus it is now possible to aim to develop control algorithms that optimize the overall value of the crops and electricity output by serving plants just the appropriate amount of light, heat and water it needs, while being in outdoor conditions. This kind of optimization is dynamic, because the growth of a plant depends on the flows of light, heat and water it receives (past, present and future, weather forecasts may influence decisions). This optimization is also stochastic, because it depends on the weather (sunlight, temperature and rain) and the energy prices on the markets that both are very unpredictable.

One of the specificity of this problem is that it requires not only to integrate stochastic variables but also to integrate qualitative performance indicators (the taste of a glass of wine cannot be scientifically measured and takes a few months after harvest) to be maximized. This implies a long term project, as it requires a minimum of a full year to test the advantages of this process on the crops. This minimum duration for experiments is even longer for wine production, as the vine must be planted after the panels are installed and vine do not optimally produce before a few years of adaptation to the new land. These long periods require to work a lot on simulation before going on the field, to make sure that what is experimented is efficiently defined.

2.4 ARTELYS (Pierre Girardeau)

Optimization problems are stochastic, just as all models are wrong. Engine outages occur. Measurement instruments get measures wrong. Demand estimates are anything but exact.

This being said, the more time passes, the more information people learn on their system. Still, a decision in such a stochastic environment as reality is always some sort of educated guess. But, as forecasts get better with time, it is just common sense that using this additional information to think again our decisions generally has some added value.

While this "perpetual decision questioning" process clearly makes sense economically, the share of optimization issues that are dealt with using a stochastic optimization modeling framework is very limited. Now, why would that be?

In the energy sector, utilities have been using stochastic optimization techniques for a long time. Considering the amount of uncertainties that can affect energy systems (commodity prices, energy demands, water inflows etc.), optimizing generation management and investments using an explicit modeling of randomness is almost impossible to avoid.

However, in our opinion, the variability of inflows and demands is not sufficient to justify the early and intensive use of stochastic optimization in this area. The economic impact of generation management and investment decisions is major, either for a large company or even for a State. For this reason, the approximations that would follow from a deterministic-only decision process would lead to discrepancies in the evaluation and optimization of energy assets that could sum up to billions of euros.

In the field of logistics, on the other hand, where optimization techniques have been intensively used for quite a while, deterministic models are often preferred. In this area, just as for energy systems, the amount of randomness in the optimization model is huge (outages, demands, delays, to name a few). However, it is not considered a major issue. Though randomness cannot be completely neglected, these uncertainties generally cannot change drastically the profits or losses of a large company.

Hence it can be argued that, up to now, stochastic optimization is only used when it is really needed, that is when it is not really possible to deal with the issue in another way. Why such a situation?

Firstly, stochastic optimization is usually very expensive. Of course, high performance computers are now quite easily available and very efficient algorithms have been developed. But it is not the main issue. That would be data. For stochastic optimization to be of any use, one needs a very important amount of data and of forecasting models that have to be structured in a right way and updated as often as possible. These forecasting models are at the heart of decision making; if they are wrong, decisions will be based on a biased vision of the future... In other words, this information has to be very reliable and trusted by users. If not, simplified models, as deterministic models, should clearly be preferred.

Secondly, the intelligibility in decision management processes is a very key issue. To be useful, the results of an optimization model have to be reliable and understandable. For the optimization consultants at Artelys, for instance, this is an everyday challenge, even when deterministic models are used. In that regard, the use of a stochastic model makes the consultancy job even more complex: this complexity has to be sufficiently clear so that the results and their limits can be perfectly understood by decision makers. This comes with a significant price. Making a stochastic optimization model, for a particular application, completely understandable and reliable is way more expensive than it is for deterministic models.

Now, like any promising technology that spreads out, more and more decision makers start being convinced of the interest of stochastic optimization for real-world problems. The rise of the "big data" ecosystem allows for more and more reliable statistical models of random processes and evaluations of strategies on a larger number of scenarios. With cloud computing, it is now easier for small companies to get an access to high-performance computers.

For years now, Artelys has chosen to be an actor in the larger adoption of stochastic optimization modeling.

First, statisticians at Artelys make use of big data to create stochastic models for energy prices, intermittent productions, and power demands. These models have already been used

in major studies for utilities, transmission system operators, regulators, ministries.

Then, our numerical optimization experts develop high performance algorithms to efficiently solve optimization problems with particular structures, among which stochastic problems for energy management and investment planning. The POST project (started in 2013), financed by the French Agency for Environment (ADEME) and leaded by Artelys, includes a partnership with INRIA on high-performance computing. It aims at providing reliable solutions to the question of energy investment planning on a European scale.

Finally, Artelys IT engineers have been working for almost ten years on creating the best graphical ways to make our clients understand the advantages and limits of stochastic models, through the development of our software suite Artelys Crystal. Artelys distributes pieces of software with comprehensive graphical user interfaces, each one dedicated to a specific application: long-term investment planning, maintenance planning, short-term power scheduling, local energy policies evaluation, risk evaluation for energy market participants.

Optimization problems are stochastic.

2.5 RTE (Patrick Panciatici)

2.5.1 Needs from a grid operator perspective

The electrical grids and their management become more and more complex. This state of affairs has different causes that will not disappear in the near future.

In Europe, the first reason is the massive integration of renewable but generally intermittent generation in the system. Power flows in the grid are created by differences in the location of sinks and sources. With a significant amount of intermittent generation, the predictability of the sources (location and amount of power injections) decreases and affects the predictability of the flows. Furthermore, some of these new power plants could be small units (e.g. PV) connected to the distribution grid, changing the distribution grid into an active system. Moreover, Transmission System Operators (TSOs) have a poor observability of these power injections and have no control at all over them. Another factor is the inconsistency between the relatively short time to build new wind farms (2 or 3 years) and the time to go through all administrative procedures to build new lines (more than 5 years everywhere in Europe). In Europe, the best locations for wind farms are mostly along the coasts and offshore, while for photo-voltaic generation they are in the south of Europe. Since these locations do not generally match those of the large load centers, a transmission network is required and this network will have to cope with the variability of the flows induced by the stochastic nature of the novel generation subsystems.

The second main reason is that it is more difficult than ever to build new overhead lines because of low public acceptance and "Not In My BackYard" (NIMBY) attitude. People are more and more afraid of hypothetical electromagnetic effects or just don't like to see big towers in the landscape and in particular in protected areas which are more and more numerous around Europe. It is very difficult to explain the need for new power lines to people who already have access to electricity at a reasonable price and with high reliability. An

increase in the European Social Welfare with a positive feedback for the European economy and hopefully for all European citizens is a concept that is too theoretical compared to the negative local impact. Alternative solutions are technically complex, costly and need more time to be deployed.

The third reason is linked to the setup of electricity markets crossing the administrative and historical borders. Generators, retailers and consumers view the transmission system as a public resource to which they should have unlimited access. This approach has the desirable effect of pushing the system towards a maximization of the social welfare and an optimal utilization of the assets. However, this optimization is constrained by security considerations because wide-spread service interruptions spanning over long periods of time are unacceptable in our modern societies due to their huge economic and social costs. Since TSOs are responsible for maintaining the reliability of the electric power system, they must therefore define the operating limits that must be respected. As in any constrained optimization problem, the optimal solution towards which the market evolves tends to be limited by these security constraints. The stakeholders therefore perceive reliability management by the TSOs as constraining their activities and reducing the European Social Welfare rather than as enablers of this large physical market place, as it would be the case if the grid were a copper plate. The transparent definition and the precise assessment of the distance to these limits thus become more and more critical.

The last reason is that the aging of grid assets needs increasing attention. A significant part of the European grids' assets are more than 50 years old. Asset management, and maintenance in systems that can't be stopped, are extremely challenging and need to be precisely anticipated when large numbers of assets are approaching simultaneously the end of their expected life times.

To maintain the security of the supply in this context, TSOs have to change the architecture of the system by considering the following technologies:

- Long distance HVAC underground cables with large reactive compensators;
- HVDC underground cables in parallel with the AC grid with smart controls of AC/DC converters;
- And, ultimately, HVDC grids, first to connect efficiently offshore wind farms and then to provide cheaper interconnections between distant areas.

Meanwhile, TSOs will try to optimize the existing systems by adding more and more special devices such as Phase Shifting Transformers, Static VAr Compensators and advanced controls and protection schemes, taking also advantage of the flexibility provided by HVDC links embedded in AC grids. At the same time, demand response or dispersed storage could offer new ways to control the system, even if business models and costs are still questionable. But in any case, this flexibility will require a rethinking of historical operating practices where grid operators made the assumption that the load is an uncontrollable exogenous stochastic variable.

We have heard so often in conferences, seminars and workshops, that the power grid will soon be operated very near to its limits, so that this statement has become a cliché. This

cliché is now a reality. To be more precise, it is no longer possible to respect the classical preventive $N-1$ security standards during all hours in a year. The system is indeed no longer able to survive all single faults without post-fault actions, i.e. corrective controls. More and more corrective control strategies are hence elaborated and prepared to maintain the security of the system. The number of hours during which the system requires corrective actions to be secure is increasing, and that seems to be a natural trend associated with the massive integration of intermittent generation. More and more local or centralized Special Protection Schemes (SPS)/Remedial Actions Schemes(RAS) are deployed to implement automatically some of these corrective actions based on advanced measurement devices (Phasor Measurement Units, Dynamic Line Ratings, ...) and high bandwidth communication networks.

Grid operators have to manage an extremely complex decision making process in order to ensure the reliability and quality of supply at minimal cost over different time horizons. For the sake of clarity, while not aiming at being exhaustive, the following problems need to be dressed by the grid operators:

- Long term (10-20 years): planning stage
  - where to build new power lines? which technology? which capacity?
- Mid term (2-5 years):
  - installation of control devices: substation design, var/reactive support, PSTs, replacement of conductors, SPS/RAS design;
  - asset management and maintenance: which equipment to upgrade, to replace, to repair and when?
- Short term (monthly-weekly):
  - outage management, must-run generators, preparation of corrective actions, required margins.
- Real Time (two days ahead to real time):
  - interaction with energy markets: definition of grid capacities;
  - selection of substation's topology, settings of SPS/RAS, adjustment of generating units.

In all these contexts, the grid operators want to make "optimal" decisions over these different time horizons, even if some decision making processes are currently not formalized mathematically as optimization problems but are rather based solely on knowledge of experts. However, as complexity increases, decision support tools become mandatory to help these experts to make their decisions:

- For the long term planning, there is hyper uncertainty associated with the implementation of energy transition policies and long term market behavior (Priority to renewable

energies, Demand Growth in context of efficiency promotion, Technology Costs: electrical batteries, Demand Response, EV, Distributed Generation, Carbon Tax, Fuel Costs) and Grid operators have to make robust decisions based on multi-scenario grid planning.

- In all these processes, the increasing level of uncertainty associated to wind and solar power must be taken into account, hence pushing towards the use of probabilistic methods.

- Operation nearest to the limits requires an accurate modeling of all pieces of equipment, of the corrective actions and of the dynamic behaviors, so as to allow an accurate enough assessment of security margins. Moreover, the active constraints could be related to voltage/reactive or stability issues and not only to thermal limits.

Grid operators must ensure an adequate consistency between these decision making processes. They are in fact multistage decision making processes considering all the different time horizons. At the planning stage, they have to consider the decisions which could be made in lower level problems: asset management and operation and the same between asset management and operation. The modeling of these lower level problems seems very challenging when these lower level problems become more complex. Approximations are required and relevant "proxies" must be found for this modeling.

In this paper we defend the idea that in order to address all these different questions, it is valuable to explicitly formulate them as optimization problems. Most of these problems, once stated, are hard to solve exactly. On the other hand significant progress has been collected in the recent years both in computational and in mathematical respects. Our goal is to highlight the main avenues of progress in these respects and explain how they can be leveraged for improving power systems management. In the next subsection, we present a taxonomy of optimization problems associated with the practical needs of power systems.

2.5.2 Taxonomy of optimization problems

The objective of this subsection is to dig into the different ways one can formulate and model optimization problems to be addressed in power systems (with a focus on system-wide problems, as those addressed by TSOs).

Modeling the optimization problem from a formal viewpoint

Based on examples, we discuss the intrinsic nature of the different optimization problems, by distinguishing different possible formulations.

The general formulation can be summarized as a multistage decision making problem under uncertainty. But the formulation of this very general problem depends on the different time horizons and the type of decisions to make.

We can divide these decisions in three classes, illustrated here by an analogy with IT systems.

1. Decisions changing the structure of the system (developing the hardware)
2. Decisions changing policies or control/protection schemes (developing the software)
3. Decisions modifying the operating points of the system (selecting input data to run the software on the hardware)

We focus our discussion on the first problem which is the most challenging. This problem ideally requires the modeling of all the aspects of power systems: from possible long term energy policies to system operation using realistic modeling of the physical system and expectations of the grid users.

The decisions related to the structure of the system ("hardware") take long time to be implemented, they have to go through long permitting processes and need quite long construction times. They are investment decisions. The objective is to optimize the associated capital expenditures (capex) by comparing them to future operational expenditures (opex) saving. The time frame is varying from around ten years for the construction a new power line to one year for changing conductors on existing power lines. A part of the problem is to choose the relevant mixture of technological options: ac overhead power lines, ac underground cables, hvdc links, phase shifter transformers (PSTs), new conductors for existing power lines, new reactive compensation devices, ...

The problem can be formulated as stochastic dynamic programming problem. For long term expansion planning (as illustrated in Figure 2.1), scenario-based approaches seem the most attractive formulations in order to ensure some level of robustness as proposed in [56].

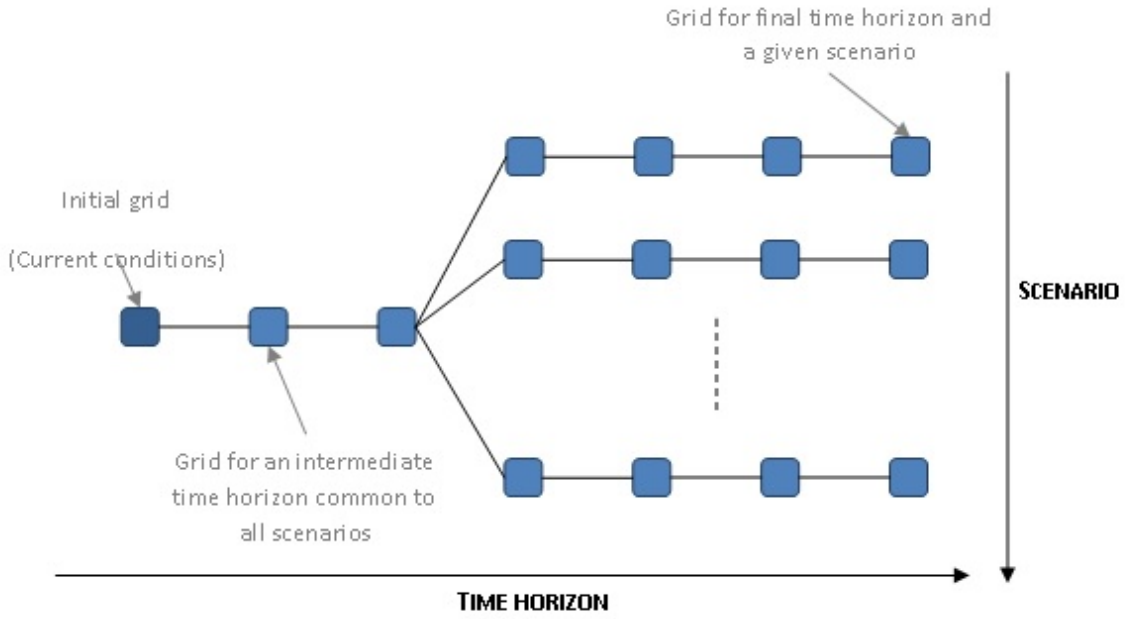


Figure 2.1: Expansion Planning

How to define reliability criteria and how to implement them, are key questions in these optimization problems. "Energy Not Served" or "Loss of Load" are generally used. An "artificial" monetization is performed and estimated costs are associated to these indexes. These large costs are simply added to operational expenditures. In stochastic formulations, generally only expected values are minimized without any cap on the maximum risk. This could be questionable and chance constraint programming or robust optimization could offer more relevant solutions. We could imagine that a generalization of Demand Response could change dramatically the definition of reliability and the foundations of power system design, pushing to less "hardware" and more "software" solutions as anticipated very optimistically in 1978 by F. C. Schweppe [61].

A review of the current formulation and associated optimization problem is mandatory as proposed for example in the two on-going European projects: e-HIGHWAY2050 (Modular Development Plan of the Pan-European Transmission System 2050) and GARPUR (Generally Accepted Reliability Principle with Uncertainty modelling and through probabilistic Risk assessment). In this global optimization problem, the sub problem on selection of relevant technological options leads to a combinatorial optimization very similar to a "knapsack" problem [1], increasing even more the complexity.

We can identify three different dimensions: spatial, temporal and stochastic. The spatial complexity is increasing: "more and more the electrical phenomena don't stop at administrative borders". We have to consider systems very extended (Pan-European Transmission System, Eastern or Western Interconnection in US, ...) and at the same time local active distribution grids. Time constants range from milliseconds to several years, leading to temporal complexity. Uncontrollable load and renewable energy sources implies to take into account more than ever stochastic behaviors. Considering spatial, temporal and stochastic complexity all together is still out of reach. Trade-offs must be made to take into account at most two of them in details at the same time and using approximation for third one.

The appropriate modeling of uncertainties is also a key factor to find realistic optimal solutions. The spatial and temporal correlations between these uncertainties must be taken into account not to be too optimistic or too pessimistic. This pushes towards probabilistic methods and risk based approaches. When the probabilistic properties of the uncertainties are only partially known, generalized semi-infinite programming seems an appealing method proposing to find robust solutions when the uncertainties live in a defined domain:

$$\min_{x \in X} f(x)$$
$$\text{subject to: } \forall \delta \in \Delta : g(x, \delta) \leq 0,$$
where x are the decision variables and δ the uncertainties.

The objective function is related to satisfaction of the grid users (consumers and suppliers): maximization of social welfare. We need to estimate the expectations and the behaviors of the grid users. For long term decisions, it seems reasonable to simulate a "perfect market"

[1] the knapsack problem: given a set of items, each with a mass and a value, determine those to include in a collection so that the total weight is less than or equal to a given limit and the total value is as large as possible.

leading to a global rational behavior minimizing the costs. For more short term decisions, it could be important to simulate the actual behavior of the market players and the imperfect market design. These estimations could be formulated as optimization problems based on game theory finding Nash equilibrium[2].

In practice, we could have to formulate multi-objective optimization problems which are generally transformed in a single optimization problem using a weighted sum of the different objective functions

$$\min(w_1.f_1(x) + w_2.f_2(x) + ... + w_n.f_n(x)).$$

(For example, we want to minimize the production costs and the amount of emitted CO_2). Finding the associated weighting factors could be difficult and questionable. A more rational approach should be to formulate a true multi-objective function

$$\min(f_1(x), f_2(x), ..., f_n(x)).$$

But for a nontrivial multi-objective optimization problem, there does not exist a single solution that simultaneously optimizes each objective. In that case, the objective functions are said to be conflicting, and there exists a (possibly infinite number of) Pareto optimal solutions. A solution is called nondominated, Pareto optimal, Pareto efficient or noninferior, if none of the objective functions can be improved in value without degrading some of the other objective values. This leads to complex optimization problems which could be solved using meta-heuristics methods.

We could see that power system management could lead to a large diversity of optimization problems. The proper formulation of each problem has to be well thought out before searching for computational solutions.

Modeling the physics of the power system

The objective of this subsection is to analyze the physics and technological constraints arising from the power system, and explain their implications in terms of the nature of the above formulated optimization problems.

The quality of physical modeling of power systems used in optimization problems is essential in order to make "optimal" decisions. Solving optimization problems with a high accuracy based on not realistic enough modeling is useless. We need to find the right balance between realism and complexity. The usage of static and deterministic modeling using a linearization of the associated mathematical formulations should be questioned in the new context presented in the introduction.

A significant number of controls in electrical grids are discrete: switch on/off of breakers, switch on/off capacitor or reactor banks, tap changers on transformers, generating units producing with non zero minimal active power when they are started. These controls become integer variables in optimization problems and their treatments require a special attention;

[2]Nash equilibrium is a solution concept of a non-cooperative game involving two or more players, in which each player is assumed to know the equilibrium strategies of the other players, and no player has anything to gain by changing only their own strategy

with a naive relaxation (round off strategy) is not always possible even to find feasible solutions.

Some controls and protection schemes implemented in local or centralized SPS/RAS are not event-based but measurement-based. They acts conditionally when a measurement or a set of measurements don't fulfill a given rule (for example, when a measurement value is beyond a given limit). This kind of behavior must be taken into account in optimization problem. This leads to conditional corrective actions and the modeling of this type of hybrid system (continuous and discrete) requires binary variable [7].

Some active constraints in power systems could be more and more related to stability and the system dynamic behavior. The ultimate solution should be to use a DAE-constrained optimization formulation but a reasonable first step could be to use "proxies" to hide this complexity. The idea is to learn using Monte-Carlo simulation, rules which ensure that the study case has a very low probability to be prone to stability issues [76] and to introduce these rules in static optimization problems.

Power system planning and operation raises many important decision making problems, which can generally be stated as large-scale, non-linear, mixed-integer continuous, non-convex, stochastic and/or robust optimization problems.

In the last years, many progresses have been made in the theory and implementation of optimization algorithms, driven by research in applied mathematics and by multitudinous opportunities of application. The combination of these novel ideas to improve the state-of-the-art of power systems optimization is an important direction of future work.

On the other hand, low cost information technology (HPC and Big Data) as well as progresses in machine learning and randomized algorithms offer other enabling approaches to apply optimization techniques in power systems.

We suggest that the research community should further focus on the proper formulation of power system optimization problems with the help of power system experts, and develop more intensively fruitful collaborations with researchers in applied mathematics and computer science to determine the most effective solution strategies for these problems.

At the same time, we think that more systematic investments in a more effective use of modern information technologies, especially in the context of high-performance computing and massive data exploitation should be made by the power systems industry.

2.6 Électricité de France R&D (Marc Porcheron)

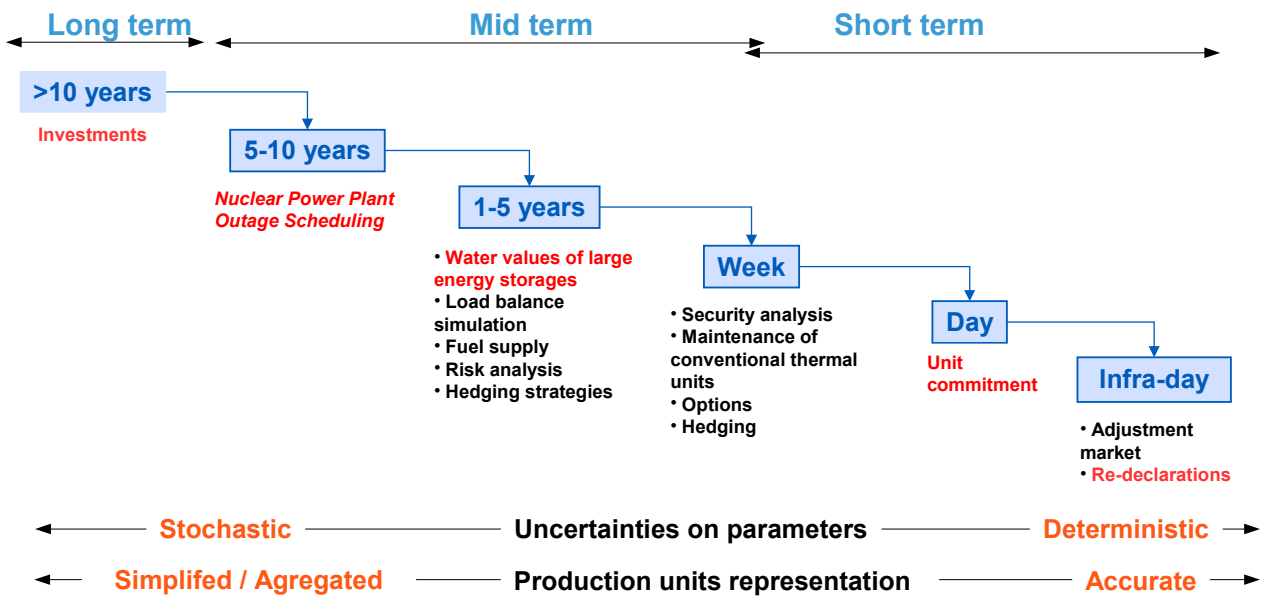
With about 40 million customers around the world, and a worldwide workforce of near 160,000 employees, EDF is one of the world's largest electricity suppliers. This position puts EDF at the forefront of the challenge of responding to growing energy demand while tackling climate change and managing resource depletion.

The transition toward a carbon-free energy economy means stepping up to new problems, including: managing intermittency of renewable energies as best as possible, incorporating new uses of electricity by optimizing production methods and grids, devising energy management systems on a local scale, and on a larger one, developing grid infrastructures and

optimizing electricity flow in Europe.
In this case study, we focus on the first problem listed above, namely the managing of intermittency of renewable energies, taking as an illustration one of the important optimization problems tackled by EDF: the scheduling of nuclear reactor outages.

2.6.1 EDF energy management process

The following figure illustrates the general energy management process implemented by EDF in order to optimize its production portfolio. Optimization is performed at different time horizons, the output of each problem serving as an input for the ones dealing with closer time horizons:



Up to now, the conventional picture was as follows: problems on far time horizons were of course heavily stochastic due to the uncertainty that affects their parameters, but they

did not need an accurate representation of the production unit operations; conversely, short term problems were usually deterministic while requiring a very precise representation of the production constraints.

The current evolutions of the energy system listed in introduction dramatically changes this picture, invalidating some usual approximations:

- Deterministic view of the problems in no longer acceptable, even on the short term horizons;
- Relaxations of some unit operating constraints usually performed are no longer possible, even on the long term horizons;
- Decomposition schemes usually implemented may be obsolete (Eg. does a "price decomposition" scheme still make sense in a totally decentralized production/ distribution/ consumption system?).

To sum up, we need more accurate, more stochastic models, and new resolution methods to cope with them.

2.6.2 Impacts of renewable energies penetration on EDF optimization models: an example

To illustrate the previous point, let us take as the an example one of the important optimization problems tackled by EDF: the scheduling of nuclear reactor outages. Briefly stated, it consists in optimizing each month the outage schedule and the refuel quantities of the 58 nuclear reactors of EDF on a time horizon of 5 to 10 years.

Due to this far time horizon, this problem is of course heavily stochastic. Random parameters include:

- *Demand* to satisfy at every time step;
- *Price and volume* of buying/selling on the spot market;
- *Outages duration* of the nuclear reactors;
- *Maximum and minimum power* of the nuclear reactors, affected by incidents between the outages;
- *Current stock* of the nuclear reactors;
- *Maximum power* of the conventional thermal units, affected by hazardous faults.

The objective is to minimize the nuclear reactor refueling cost plus the expectation of the production cost on this stochastic space, while satisfying:

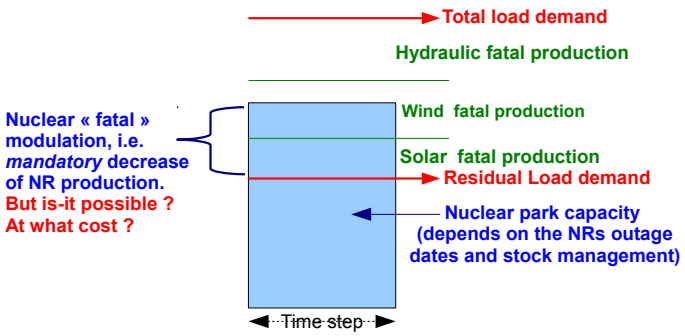
- electricity demand for every time step;

- nuclear reactors and conventional production unit operating constraints;
- scheduling and resources constraints to be fulfilled by the nuclear reactors outages.

This is a very big, stochastic, combinatorial, non-linear problem, which, with the help of numerous auxiliary binary variables, can be modeled as a *Mix Integer Linear Problem* (*MILP*). Due to its huge size, this *MILP* cannot be solved directly. Common simplifications performed to tackle it consist in:

- solving the problem on a average random event scenario;
- using a extended time step, typically covering a week.

Let us now show how the penetration of *Renewable Energies* in the energy system tends to invalidate theses approximations.
Hydraulic, wind and solar productions are *random, intermittent* productions a significant part of which can be viewed as *fatal*, i.e. for which *no storage is available*. On some occurrences, this fatal available energy *lowers* the load demand that nuclear reactors must fulfilled, consequently increasing the so-called *fatal modulation* of the nuclear park:



But due to operational constraints, this decrease of nuclear production is not always possible without stopping reactors currently connected to the grid. A complex and costly operation. For example, stopping a nuclear reactor in order to adapt the load balance during a 4 hours wind peak, will oblige to keep this reactor off-line at least for the whole day, leading to a costly call to some conventional thermal units (e.g. gaz, coal, oil) in order to replace its production.
It is worth noting that, because of the important nuclear capacity of the French park, this problem of managing the nuclear fatal modulation did exist before the expansion of wind and solar productions. But current studies tend to demonstrate that the expected development of these energies in the next few years will make these situations more frequent and, consequently:

1. we need to consider such scenarios in the problem, and thus *an average view of the random demand is no longer acceptable*;

2. we need to consider the accurate operating constraints that should limit the nuclear reactor modulation, and thus *a weekly time step is no longer acceptable.*

In this perspective, we are currently investigating with our academic partners, among them the RealOpt team of INRIA and the LIPN team of Université Paris Nord, new optimization approaches to solve this problem by combining advanced decomposition schemes:

- Dantzig-Wolfe decomposition to compute by column generation solutions for the "robust" schedule of outages and reload quantities, viewed as first-stage decisions;

- Benders decomposition to generate cuts eliminating fist-stage solutions on the basis of production constraints, productions being viewed as "recourse" second-stage decisions, depending on the scenarios.

2.6.3 Conclusions

In this short contribution, we tried to illustrate on a very concrete example some impacts of the tremendous changes lived by the energy system on the optimization models currently used by EDF to manage its generation resources. Beyond this peculiar example, it seems clear that these evolutions, from the penetration of renewable energies to the development of smart grids dealing with the division of energy management systems on different scales, oblige us to reconsider some assumptions currently made in our models, and to search for new optimizations methods. This appears to be particularly true in the field of discrete optimization dedicated to large stochastic combinatorial problems.

2.7 Alstom Grid (Laurent Schmitt)

Alstom Grid is a technology leader in the areas of Market Management, Energy Management, Distribution Management and Demand Response Management supplying control room

software platform to several large energy utilities in the world such as PJM, PG&E, Duke Energy in the US or Rte, Erdf, Energinet.dk and Stedin in Europe as well as PowerGrid in India and State Grid in China.

Operational dispatch Optimization is a core element of several of platforms, evolving from traditional applications in the areas of Unit Commitment, economic dispatch, energy and grid ancillary service market clearing, optimal power flow calculations, as well as Market participant resource portfolio optimization within intraday and realtime balancing mechanisms. These technologies have historically been derived from Lagrangian relaxation methods 15 years ago towards Mixed Integer programing and robust optimization in more recent years taking advantage of significant improvements seen in mathematical modelling languages and underlying calculation engines as well as computing.

SmartGrids have opened new areas of development new optimization of Distributed Energy Ressource portfolios - Demand, Storage, Electrical Vehicule and Distributed Generation while raising new expectations in term of renewable and load forecasting. Over 20 demonstration projects are currently implemented across the world evolving concepts from historical central optimization schemes towards layered optimization architectures leveraging new agent based technology at the lowest levels in the energy distribution systems. Key projects in that space are Reflexe and Nicegrid in France with Dalkia and Erdf as well as Duke DERMS and Nevada Energy Demand Response projects in the US.

Chapter 3

Uncertainty in Decision-Making Can Be Handled in Many Ways

Since decades, the energy industrial sector uses optimization methods to solve such problems such as unit commitment, dams management, optimal power flow, etc. As new problems emerge, discussed in Chapter 1, the mathematical discipline of optimization is facing new challenges, illustrated in Chapter 2. In this Chapter 3, we try hard to show how such challenges change the way optimization problems are framed (whereas Chapter 4 will be devoted to outlining resolution methods, especially in the realm of stochastic optimization).

In §3.1, we work out toy examples in energy management. In the process of incorporating uncertainties into simple optimization problems, we hope the reader will get a good feel for the kind of mathematical machinery needed. In §3.2, we outline the ingredients for framing optimization problems under uncertainty, so that we can develop mathematical formulations in §3.3. Finally, as our exposition is till now restricted to the so-called risk neutral approach, we discuss in §3.4 a variety of ways to handle risk attitudes in optimization.

3.1 Working Out Toy Examples in Energy Management

Here, we work out several toy examples in energy management to illustrate how uncertainties can be incorporated into simple optimization problems. We put the emphasis on the two main new issues in *stochastic control* in comparison with *deterministic* control: *risk* attitudes and online *information*.

3.1.1 Stochastic economic dispatch

We start by working out a first simple example of economic dispatch as a cost-minimization problem under supply-demand balance.

Problem statement

We consider two energy *production* units

- a "cheap" limited one which can produce a quantity q_0, with $0 \leq q_0 \leq q_0^\sharp$, at cost $c_0 q_0$
- an "expensive" unlimited one which can produce quantity q_1, with $0 \leq q_1$, at cost $c_1 q_1$, with $c_1 > c_0$

On the *consumption* side, the demand is $D \geq 0$.

We express the *supply-demand balance objective* as ensuring at least the demand, that is

$$q_0 + q_1 \geq D \; . \tag{3.1}$$

This objective is to be achieved at least cost, so that the *optimization* problems is:

$$\min_{q_0, q_1} \underbrace{c_0 q_0 + c_1 q_1}_{\text{total costs}} \; . \tag{3.2}$$

When the demand D is deterministic, the optimization problem is well posed

When the demand D is deterministic, D is a single known number, and we minimize

$$\min_{q_0, q_1} c_0 q_0 + c_1 q_1 \text{ , under the constraints } \begin{array}{rl} 0 & \leq q_0 \leq q_0^\sharp \\ 0 & \leq q_1 \\ D & \leq q_0 + q_1 \end{array} \tag{3.3}$$

The solution is

$$q_0^\star = \min\{q_0^\sharp, D\} \, , \quad q_1^\star = [D - q_0^\sharp]_+ \, , \tag{3.4}$$

that is,

- if the demand D is below the capacity $q_0^\sharp$ of the "cheap" energy source, you only appeal to this source:
$$D \leq q_0^\sharp \Rightarrow q_0^\star = D \, , \quad q_1^\star = 0 \, , \tag{3.5}$$
- if the demand D is above the capacity $q_0^\sharp$ of the "cheap" energy source, you have to have recourse to the "expensive" source:
$$D > q_0^\sharp \Rightarrow q_0^\star = q_0^\sharp \, , \quad q_1^\star = D - q_0^\sharp \; . \tag{3.6}$$

Now, what happens when the demand D is no longer deterministic?

From now on, we suppose that the demand D is a random variable $D : \Omega \to \mathbb{R}_+$, with mathematical expectation $\mathbb{E}(D) = \overline{D}$.

What happens if we replace the uncertain value D of the demand by its mean $\overline{D}$ in the deterministic solution?

To forge a "solution" (to which problem? we will come to that later), one can propose the "deterministic solution"

$$q_0^{(\overline{D})} = \min\{q_0^\sharp, \overline{D}\} \ , \ \ q_1^{(\overline{D})} = [\overline{D} - q_0^\sharp]_+ \ . \tag{3.7}$$

Now, the question which arises is: how can I assure the supply-demand inequality

$$\underbrace{D(\omega)}_{\text{uncertain}} \ \leq \ \underbrace{q_0 + q_1}_{\text{deterministic}} \ \ , \ \forall \omega \in \Omega \ , \tag{3.8}$$

using the deterministic solution $(q_0^{(\overline{D})}, q_1^{(\overline{D})})$? On the one hand, Equation (3.8) corresponds to an infinite number of constraints, which are all satisfied if the condition $\sup_{\omega\in\Omega} D(\omega) \leq q_0^{(\overline{D})} + q_1^{(\overline{D})}$ holds true. On the other hand, Equation (3.7) implies that $\overline{D} = q_0^{(\overline{D})} + q_1^{(\overline{D})}$; we arrive at a contradiction since $\overline{D} < \sup_{\omega\in\Omega} D(\omega)$ (except in the deterministic case where $D(\omega)$ does not depend on ω). Therefore, the deterministic solution obtained by solving the problem "in mean" is not acceptable.

When the demand D is bounded above, we are able to obtain a "robust" deterministic solution

The so-called "robust approach" consists in taking into account all possible constraints in the optimization problem. In the robust optimization formulation, we thus minimize

$$\min_{q_0,q_1} c_0 q_0 + c_1 q_1 \ , \ \text{ under the constraints } \quad \begin{array}{rll} 0 & \leq q_0 \leq q_0^\sharp & \\ 0 & \leq q_1 & \\ D(\omega) & \leq q_0 + q_1 & \forall \omega \in \Omega \end{array} \tag{3.9}$$

When $D^\sharp = \sup_{\omega\in\Omega} D(\omega) < +\infty$, that is, when the demand D is bounded above, the solution to (3.9) is

$$q_0^\star = \min\{q_0^\sharp, D^\sharp\} \ , \ \ q_1^\star = [D^\sharp - q_0^\sharp]_+ \ . \tag{3.10}$$

Now, we observe that the total cost $c_0 q_0^\star + c_1 q_1^\star$ is an increasing function of the upper bound $D^\sharp$ of the demand. So, whatever tiny may be the chance that such high value materializes, the value $D^\sharp$ drives the "solution" and the total cost. Therefore, one can wonder if it is not too costly to optimize under the worst-case situation.

What happens if the problem is solved demand value by demand value?

If the problem (3.3) is solved for each possible value $d = D(\omega)$ of the random variable D, when $\omega \in \Omega$, we obtain a collection of "solutions"

$$q_0^{(d)} = \min\{q_0^\sharp, d\} \ , \ \ q_1^{(d)} = [d - q_0^\sharp]_+ \ . \tag{3.11}$$

Now, we face an *informational issue*:

- if the demand D is observed before selecting the quantities q_0 and q_1, this collection of "solutions" is optimal in many understandings;
- whereas, on the contrary, if the demand D is not observed, how can one glue together those "solutions" to cook up quantities q_0 or q_1 that do not depend upon the unknown quantities d?

When the demand D is not observed, there is no proper way to do the gluing.

To overcome the above difficulties, we turn to stochastic optimization

We now consider a new optimization problem

$$\min_{q_0,q_1} \mathbb{E}[c_0 q_0 + c_1 q_1] \ , \quad \text{under the constraints} \quad \begin{array}{ll} 0 & \leq q_0 \leq q_0^\sharp \\ 0 & \leq q_1 \\ D & \leq q_0 + q_1 \\ q_1 & \text{depends upon } D \ , \end{array} \tag{3.12}$$

where $\mathbb{E}$ denotes the mathematical expectation (in the probability space Ω where the random variable D is defined). This framing raises two issues, that are new with respect to the deterministic case:

- expliciting online information issue: the decision q_1 depends upon the random variable D;
- expliciting risk attitudes: the total costs are aggregated with respect to all possible values by taking the expectation $\mathbb{E}[c_0 q_0 + c_1 q_1]$.

Turning to stochastic optimization forces one to specify online information

Specifying in (3.12) that the decision q_1 depends upon the random variable D, whereas q_0 does not, forces to consider two stages and a so-called non-anticipativity constraint (more on that later):

- first stage: q_0 does not depend upon the random variable D;
- second stage: q_1 depends upon the random variable D.

Turning to stochastic optimization forces one to specify attitudes towards risk

Now that q_1 depends upon the random variable D, it is also a random variable, and so is the total cost $c_0 q_0 + c_1 q_1$. Therefore, we have to aggregate the total costs with respect to all possible values, and we chose to do it by taking the expectation $\mathbb{E}[c_0 q_0 + c_1 q_1]$.

Where are we now?

In the uncertain framework, two additional questions must be answered with respect to the deterministic case:

- How are the uncertainties taken into account in the payoff criterion and in the constraints? Answering this question amounts to expliciting the risk attitude of the decision-maker.
- Upon which online information are decisions made? This question revolves around the notion of "solution" of the optimization problem.

3.1.2 Aggregating a random source with storage

In this example, we consider an energy system whose behavior is studied over a discrete time span $\{0, \dots, T\}$. The system incorporates

- a fatal production of renewable energy at each time t, denoted $Q_R(t)$,
- a given demand $D(t)$ to be met at each time t,
- a market allowing to buy/sell, at time t, an amount $Q_M(t)$ of energy at a price $P_M(t)$,
- a stock of energy whose volume $S(t)$ has to remain within given bounds $[\underline{S}, \overline{S}]$; the stock device is able to produce/consum energy Q_S, at an operating cost denoted by $C(Q_S)$; consumption will be materialized by $Q_S > 0$, and production by $Q_S < 0$.

At each time t, the demand has to be satisfied:

$$Q_R(t) + Q_S(t) + Q_M(t) - D(t) = 0 \ . \tag{3.13a}$$

The behavior of the stock is given by the dynamics[1]

$$S(t+1) = S(t) - Q_S(t) \ , \tag{3.13b}$$

where the initial stock volume at time 0 is given, equal to s_0. The stock volume is subject to bounds constraints:

$$S(t) \in [\underline{S}, \overline{S}] \ . \tag{3.13c}$$

The overall cost is obtained by summing the operating costs of the stock, the costs induced by the market, and by including a term K depending of the stock volume at the final time T:[2]

$$\sum_{t=0}^{T-1} \Big(C\big(Q_S(t)\big) + P_M(t).Q_M(t) \Big) + K\big(S(T)\big) \ . \tag{3.13d}$$

[1] Recall that consumption is materialized by $Q_S > 0$, and production by $Q_S < 0$.

[2] This last term is added to ensure that the stock will not be empty at the end of the time span.

In a deterministic setting, the problem is a classical one

In absence of uncertainties, Problem (3.13) pertains to the class of so-called "optimal control problems" (deterministic). It can be efficiently solved using Pontryaguin principle and associated numerical methods. Moreover, in the case where the cost C is a piecewise linear function, Linear Programming applies. In all cases, the solution obtained by solving the problem leads to a collection, indexed by time, of decisions $\big(Q_S(t), Q_M(t)\big)$. Such a solution corresponds to the standard notion of *planning*, where a solution *only* depends on time.

In a stochastic setting, further specifications are mandatory

As already mentioned in the example presented in §3.1.1, turning to a stochastic framework forces one to specify online information and risk attitudes. For these latter, we assume that both the fatal productions $Q_R(t)$ and the market prices $P_M(t)$ are random variables, and we envisage risk by means of the mathematical expectation. Now, we concentrate on the issue of online information: the question amounts to explicit the information available at time t when elaborating the decisions $Q_S(t)$ and $Q_M(t)$.

The most common situation is the one where uncertainties are observed and recorded. Then, the decisions at time t depend on all past observations, namely $\big(Q_R(\tau), P_M(\tau)\big)$ for all $\tau \leq t$. The decisions are thus not only indexed by time (as in the deterministic case), but are moreover functions of the past observations, and are thus random variables (as the fatal productions $Q_R(t)$ and the market prices $P_M(t)$ are). The notion of planning does not make sense in this setting, the optimal decisions depending on both time *and* information.

Note that it would not be reasonable, either to assume a complete knowledge of the uncertainties, or the total absence of information. In the first case, decision at time t is allowed to depend on values of the uncertainties in the future, that is to say the decision maker is clairvoyant. In the second case, the absence of online information leads to decisions indexed only by time (as in the deterministic setting), but these decisions have to fulfill the equilibrium relation (3.13a) for all possible values of $Q_R(t)$, which is usually impossible.

3.1.3 Managing a smart grid network

The smart grid vocable evokes networks. Electricity networks raise important issues in what concerns phase, frequency and voltage; such issues are addressed by electrical engineers for decades and are not our concern. Here, we go on focusing on production/consumption, and we highlight issues in optimization when decision-makers are attached to locations on a network, and do not necessarily all share the same information. This example is an opportunity to touch *team theory* — where different players with different informations jointly contribute to optimize the same criterion — and *game theory*, where each player optimizes his/her own criterion.

We consider the interaction between an active producer and an active consumer:

- the producer uses a control variable u_P to drive production $q_P(u_P)$, for a cost $j_P(u_P)$;

- the consumer uses a control variable u_C to monitor consumption $q_C(u_C)$, for a cost $j_C(u_C)$.

The goal is that the production meets the consumption, that is,

$$q_C(u_C) - q_P(u_P) = 0 \ . \tag{3.14}$$

Deterministic setting

We first assume a deterministic formulation. A solution can be obtained by solving the optimization problem

$$\min_{u_P,u_C} j_P(u_P) + j_C(u_C) \qquad \text{under the constraint} \qquad q_C(u_C) - q_P(u_P) = 0 \ ,$$

provided that the constraint(3.14) is met for at least one couple of controls. Note that this problem can be tackled by the duality theory: introducing the multiplier (or price) π associated to the constraint, one solves the two subproblems

$$\min_{u_P} j_P(u_P) - \pi \cdot q_P(u_P) \text{ and } \min_{u_C} j_C(u_C) + \pi \cdot q_C(u_C) \ ,$$

and iteratively modifies the price π until $q_C(u_C) - q_P(u_P) = 0$ (Walras groping).

Centralized stochastic setting

We now assume that uncertainties affect both the producer and consumer. We denote by ω_P (resp. ω_C) the uncertainty for the producer (resp. consumer), and we turn to stochastic optimization. As explained in the toy example in §3.1.1, we have to deal with the issues of online information and risk attitude:

- we assume that each decisions u_P and u_C may depend on both the uncertainties ω_P and ω_C, and we denote by Ω the set of all possible value $\omega = (\omega_P, \omega_C)$ of the uncertainty;
- the costs are aggregated by taking the expectation with respect to a probability measure $\mathbb{P}$ (over a proper σ-algebra);
- the equilibrium constraint (3.14) has to be met in the almost-sure sense (denoted $\mathbb{P}$-a.s.), that is, for all possible ω expect on a null-measure subset of Ω.

The problem can be tackled by solving the stochastic optimisation problem

$$\min_{u_P,u_C} \mathbb{E}\Big[j_P(u_P,\omega_P) + j_C(u_C,\omega_C)\Big] \qquad \text{subject to} \qquad \begin{array}{l} q_C(u_C,\omega_C) - q_P(u_P,\omega_P) = 0 \ \ \mathbb{P}\text{-a.s.} \ , \\ u_P \text{ depends upon } (\omega_P,\omega_C) \ , \\ u_C \text{ depends upon } (\omega_P,\omega_C) \ . \end{array}$$

In such a setting, the set of admissible controls[3] is non empty provided that, for any (ω_P, ω_C), there exists a (u_P, u_C) such that the constraint is met.

[3]that is, the set of pairs (u_C, u_P) satisfying all the constraints

Again, a solution can be obtained by using the duality theory. But, in the present situation, the price π becomes a random variable, that is, a function of (ω_P, ω_C). The computation of this random price is thus a much more delicate issue than in the deterministic case: in some sense, one has to perform the Walras groping for each possible value of the uncertainty!

Decentralized stochastic setting

We now consider a decentralized situation, where the producer (resp. the consumer) has no direct information about the uncertainty faced by the consumer (resp. the producer). In order to connect the producer and the consumer, we assume that a central planner is able to influence the behavior of these two agents by sending them a signal. For the sake of simplicity, we assume that this signal is a price denoted by π. The objective of the central planner is to elaborate a price signal π such that the production q_P is equal to the demand q_C.

In this decentralized framework, the producer elaborates a decision u_P based on the knowledge of both the local uncertainty ω_P and the price signal π sent by the central planner, e.g. by solving the optimization problem

$$\min_{u_P} \mathbb{E}\Big[j_P(u_P, \omega_P) - \pi \cdot q_P(u_P, \omega_P)\Big] \qquad \text{subject to} \qquad u_P \text{ depends upon } (\omega_P, \pi)\ .$$

In the same way, the consumer solves

$$\min_{u_C} \mathbb{E}\Big[j_C(u_C, \omega_C) + \pi \cdot q_C(u_C, \omega_C)\Big] \qquad \text{subject to} \qquad u_C \text{ depends upon } (\omega_C, \pi)\ .$$

The question is now to determine what is a "good" price signal. Let us examine the two (extreme) following cases.

- If the central planner is able to produce a price signal depending on both ω_P and ω_C, we potentially fall in the previous centralized setting, and the constraint of balance between production and consumption can be satisfied in the almost-sure sense.
- If the central planner has just the ability to send a deterministic price signal π, the producer will never be informed of the uncertainty faced by the consumer and vice-versa: the balance constraint will not be met in the almost-sure sense, except for very specific situations.

With this example, we touch the question of being able to elaborate a "good" price signal, that is, a price containing the relevant information allowing to drive decentralized actors while meeting global goals. This issue is of particular interest in the smart grid framework.

3.2 Laying Out Ingredients to Frame Optimization Problems under Uncertainty

We describe the ingredients to formalize optimization problems where some data are uncertain, where decision are made at discrete time stages, and where the availability of data may differ according to time and space. In this §3.2, we try to keep the mathematical material to a modest level, whereas we develop the mathematical aspects in the next §3.3.

The main ingredients to formulate a multi-stage optimization problem under uncertainty are *discrete time*, *uncertain*, *control* and *output* variables, *scenarios*, *strategies*, *criterion, constraints*. To account for a network structure, like in smart grids, *graph theory* will also be summoned.

3.2.1 Basic variables

Basic variables: time/stage

In energy management, time may be measured in various units (decades, years, months, week, days, intra-days) over different horizons depending on the issue: years/decades for strategic planning, years for the pluri-annual production planning problem, months/years for medium term power units scheduling, weekly time steps over a one to two-year horizon for the mid-term production planning problem, day for the daily power generators scheduling, day or intra-day for the short-term operational problem (for instance, a time step of 30 minutes over one or two days).

In what follows, *times* or *stages* are denoted by the variable t, and are supposed to take values in a known set $\mathbb{T}$: $t \in \{0, 1\}$ for two-stage optimization problems; $t \in \{t_0, t_0+1, \dots, T\}$ for multi-stage optimization problems (or $t \in \{t_0, t_0 + \Delta t, \dots, t_0 + k\Delta t\}$).

Basic variables: spatial structure

Electricity flows through networks and, as we saw, the smart grid paradigm is characterized by a growing number of decentralized means of production, storage, consumption. Networks are classically represent by a *graph* $\mathcal{G} = (\mathcal{N}, \mathcal{A})$, with *nodes* $n \in \mathcal{N}$ and *arcs* $a \in \mathcal{A}$.

As with time t, uncertain and control variables can be indexed by nodes or arcs (that is, by location), like $w_{t,a}$ or $u_{t,n}$ and take values in known sets that may depend on nodes or arcs.

Basic variables: uncertainty

Typical (primitive) *uncertain* variables are meteorological (wind, sun, rain, etc.) and temperature conditions, market electricity prices, failure occurences, etc. The decision-makers have no control whatsoever on so-called uncertain variables.

Time sequences of uncertain variables form what we call *scenarios*. *Offline information*, discussed below, is prior data about uncertain variables and scenarios, and covers various

materials like probability distributions, statistical models, scenario trees, ranges of values, etc.

In what follows, *uncertain* variables at time t are denoted by w_t and are supposed to take values in a known *uncertainty set* $\mathbb{W}_t$; *scenarios* are denoted by $w(\cdot) = (w_t, t \in \mathbb{T})$, and belong to $\Omega = \prod_{t\in\mathbb{T}} \mathbb{W}_t$, the *set of scenarios*.

Basic variables: control

Typical *control* variables are water releases in a dam, on/off status of production units in unit commitment, dispatches of production units, investment in new plants, etc. To the difference of uncertain variables, decision-makers have full control upon control variables, the possible values of which can be selected at every period (or at every node of a network) within given sets.

In what follows, *control* variables at time t are denoted by u_t and are supposed to take values in a known *control set* $\mathbb{U}_t$.

Output variables

From uncertain and control variables are constructed *output* variables: states (stocks), observations, costs, indicators, etc. For a static system, outputs at time t are functions of uncertain and control variables at the same time t. For a dynamical system, outputs at time t are functions of past uncertain and control variables before time t.

State control systems are special input-output systems where an output variable is distinguished, labeled *state*, which satisfies and induction (dynamical) equation (the state at time $t+1$ is function of the state, control and uncertainty at time t). For instance, water dams stocks often play the role of state variables in dam management. However, if the water inflows process has temporal dependency, the state has to be increased to account for memory lags. In what follows, state variables at time t are denoted by x_t and are supposed to take values in a known set $\mathbb{X}_t$.

3.2.2 Basic data

In deterministic optimization, basic data takes the form of known values for parameters. When not all parameter values are known, there still is room for partial knowledge under the form of *offline information* about the uncertain variables. We coin this information *offline* because it is a global information about the uncertain variables that is available to the decision-maker when he formalizes the optimization problem: say, you know that the demand varies between bounds, and you know the bounds. Later, we discuss *online information* in relation to the notion of solution to the optimization problem: say, some control variables are allowed to depend on the value of the demand, where some are not.

Basic data: offline information

Offline information is generally about scenarios (recall that scenarios are sequences of uncertain variables). The most common ways to describe scenarios are the following:

- a scenario tree (where branching occurs at each time stage), with probabilities attached to the leafs;
- a product set (product $\prod_{t\in\mathbb{T}} \mathbb{W}_t$ of the sets $\mathbb{W}_t$ where uncertain variables take their values), equipped with a probability distribution $\mathbb{P}$ (or with a family of probability distributions) in the stochastic approach;
- a product set $\Omega = \prod_{t\in\mathbb{T}} \mathbb{W}_t$, with a subset $\overline{\Omega} \subset \Omega$ expressing which scenarios are considered and which are not (set membership restrictions) in the robust (or worst-case) approach.

For instance, weather temperatures may be modeled by a dicrete tree, by independent random variables, or by an ARMA process. Energy prices may be supposed to follow a Markov chain, prices and temperatures may be correlated, etc. Or wind may only be supposed to vary between bounds, which may depend on time (season, day/night, etc.)

3.2.3 Solution to an optimization problem

In deterministic multi-stage optimization, a solution is a sequence $(u_t, t \in \mathbb{T})$, where $u_t \in \mathbb{U}_t$. More generally, a solution is a family of controls that may be indexed by time t and by location (nodes n and arcs a).

In multi-stage optimization under uncertainty, a control at time t (and location n or a) may depend upon the *online information* available at that time (and location). Hence, a solution is a function mapping available information (about any variable) into the control set. This function may, in practice, take different forms that we discuss below. But, before that, we develop what we call online information.

Online information

Online information is information about the values of variables that the decision-maker may use when he makes a decision (control); this information may be contingent on time t and location n or a. This notion is meaningless in deterministic optimization, where all data is known to the decision-maker. We postpone a discussion on the mathematical representation of online information to §3.3.

Following [35], a decision variable is said to be *robust* if it does not depend on the uncertainty, whereas it is *flexible* if it is allowed to depend on (part of) the uncertainty. This relates to the distinction between *open and closed loop* solutions.

In a temporal setting, it is commonly assumed that, at time t, some past variables are known, but certainly not the value of any variable at times $t+1$, $t+2\ldots$ Otherwise stated, the decision maker at time t may use observations he gathered up to time t, but what will

happen after time t remains uncertain and thus cannot be observed (hence used for decision-making): this is the so-called *non-anticipativity* property, that we will develop in §3.2.4. We insist that, in deterministic optimization, the decision-maker is clairvoyant in that he knows at time t the value of any variable at times $t+1$, $t+2\ldots$ Thus, online information specifies what variables are known at time t. For instance, in the unit commitment problem, the decision whether to mobilize a unit is taken day $\mathrm{D}-1$ without knowing the exact value of the demand of day D; by contrast, the production level of a unit or of a reserve is allowed to depend on this value.

In a spatial context, information may depend on nodes and arcs of a graph. Online information is *centralized* if the decision attached to a node or to an arc can make use of the information attached to any node and arc of the graph (single decision-maker). Online information is *decentralized* if the decision attached to a node or to an arc can only make use of the local information attached to this very node or arc (multiple decision-makers). Vertically integrated utilities (monopoly) correspond to centralized information, whereas nowadays deregulated structures induce decentralized information, and the need for adequate coordination.

For instance, energy production units have local information (water inputs in a dam) that is not necessarily shared by all units, and not necessarily in the hand of a single decision-maker. On an electric grid equipped with smart devices, the demand will be measured locally and it is quite unlikely that this information will be sent to and shared by all other devices at all other locations.

Solutions as open-loop or plannings

When information is only offline, strategies are *open-loop* strategies, also called *plannings* in that, at the beginning, the decision-maker may plan all his decisions in advance (he will not reconsider them in the course of action). Planning is the good notion of solution for deterministic optimization problems, whereas it is no longer adapted when information for decision depends on when the decision is made. Indeed, online information enlarges the set of possible solutions, leaving room for improvement in optimization.

Solutions as closed-loop or strategies

When online information feeds decision, a solution becomes a function that maps available information (about any variable) into a control set. A *strategy* (or a *policy*) is a decision rule (a function) telling each agent the control to select at given time and location for a given online information.

Strategies are *closed-loop* solutions, in that they feed the control with variables of the system under consideration. In a temporal setting, closing the loop between information and decision implicitely captures a *non-anticipativity* constraint: decisions can depend upon the past (online), but cannot depend on the future. Of course, decisions can depend upon offline or a priori information about the future (like a probabilistic knowledge of the future), but they cannot anticipate or display clairvoyance. Closed-loop strategy is the proper notion of

"solution" when we turn to stochastic optimization problems; this notion is related to those of "corrective actions", "adaptive control", "wait and see", "full recourse", etc.

Solutions "on the fly"

Even if a strategy is the proper notion of solution to an optimization problem under uncertainty, it may not be adequate in practice. Indeed, in theory, a solution is a collection of functions, that is, a collection of decision rules telling the control to envisage all times and locations for all possible values corresponding to online information. This is too much because, in practice, at a given time and location, you observe the values and you would like, "on the fly", to know just which control value to apply.

Thus, a solution "on the fly" is any mathematical device (algorithm, optimization procedure) that allows to compute the optimal *value* of a control at a given time and location, knowing online information.

3.2.4 More on the non-anticipativity constraint

Here, we discuss to more extent the issue of online information, and especially that of non-anticipativity. To speak with images, when we formalize decision-making in an uncertain context, we navigate between two stumbling blocks that are "rigidity" and "wizardry".

- On the one hand, it is suboptimal to restrict oneself, as in the deterministic case, to open-loop controls depending only upon time, thereby ignoring the available information at the moment of making a decision.
- On the other hand, it is impossible to suppose that we know in advance what will happen for all times: we exclude clairvoyance as well as look-ahead solutions.

The in-between is what is called *non-anticipativity constraints*. To give it mathematical ground, we denote the uncertainties at time t by w_t, and the control by u_t. The causal chain

$$u_{t_0} \rightsquigarrow w_{t_0} \rightsquigarrow u_{t_0+1} \rightsquigarrow w_{t_0+1} \quad \dots \quad w_{t-1} \rightsquigarrow u_t \ ,$$

models the interplay between observation and decision-making, and captures the fact that the decision at time t only depends on observations prior to t, that is, on $(w_{t_0}, \dots, w_{t-1})$. We go on by developing mathematical apparatus to tackle such *dependency*.

There are two ways to express the non-anticipativity constraint

First, we outline the two abstract formalisms used to frame the non-anticipativity constraint. Later, we will be more specific.

- Functional approach. The control u_t may be looked after under the form

  $$u_t = \phi_t\big(w_{t_0}, \dots, w_{t-1}\big) \ ,$$

  where ϕ_t is a function, called policy, strategy or decision rule.

- Algebraic approach. When uncertainties are considered as random variables (measurable mappings), the above formula for u_t expresses the measurability of the control variable u_t with respect to the past uncertainties, also written as

$$\sigma(u_t) \subset \sigma\big(w_{t_0}, \dots, w_{t-1}\big) \ ,$$

that is, as an inclusion between σ-algebras.

What is a solution at time t?

Recall that, in deterministic control, the solution u_t at time t is a single value in the control space $\mathbb{U}_t$. In stochastic control, the two abstract formalisms above induce two notions of solution u_t at time t

- either as $u_t = \phi_t\big(w_{t_0}, \dots, w_{t-1}\big)$, where $\phi_t : \mathbb{W}^{t-t_0} \to \mathbb{U}_t$,
- or as $u_t : \Omega \to \mathbb{U}_t$ with measurability constraint $\sigma(u_t) \subset \sigma\big(w_{t_0}, \dots, w_{t-1}\big)$.

Now, as time t goes on, the domain of the function ϕ_t expands, and so do the conditions $\sigma(u_t) \subset \sigma\big(w_{t_0}, \dots, w_{t-1}\big)$. Therefore, for numerical reasons, the information $\big(w_{t_0}, \dots, w_{t-1}\big)$ has to be compressed or approximated.

There are two classical ways to compress information

The two abstract formalisms and the corresponding two notions of solutions lead to the two most common way to compress information. We sketch them before discussing them in more detail.

- State-based functional approach

  In the special case of the Markovian framework — where $\big(w_{t_0}, \dots, w_T\big)$ is supposed to be a white noise — there is no loss of optimality to look for solutions as

$$u_t = \psi_t \underbrace{\big(x_t\big)}_{\text{state}} \text{ where } \underbrace{x_t \in \mathbb{X}}_{\text{fixed space}} \ , \ \underbrace{x_{t+1} = F_t(x_t, u_t, w_t)}_{\text{dynamical equation}} \ , \tag{3.15}$$

- Scenario-based measurability approach

  In the special case where the product set $\prod_{t \in \mathbb{T}} \mathbb{W}_t$ of uncertainties is approximated by a finite family of scenarios $\big(w_{t_0}^s, \dots, w_T^s\big)$, $s \in S$, we pinpoint two classical approaches.

  - Solutions $q_{i,t}^s$ are indexed by $s \in S$ with the constraint that if two scenarios coincide up to time t, so must do the controls at time t:

$$\big(w_{t_0}^s, \dots, w_{t-1}^s\big) = \big(w_{t_0}^{s'}, \dots, w_{t-1}^{s'}\big) \Rightarrow q_{i,t}^s = q_{i,t}^{s'} \ . \tag{3.16}$$

  - In the case of the scenario tree approach, the scenarios $\big(w_{t_0}^s, \dots, w_T^s\big)$, $s \in S$, are organized as a tree, and controls $q_{i,t}^n$ are indexed by nodes n on the tree

More on the state-based approach

- In the expression $u_t = \psi_t(x_t)$, the mapping ψ_t can be computed in advance (that is, at initial time t_0) and evaluated at time t on the available online information at that time t
  - either exactly (for example, by dynamic programming),
  - or approximately (for example, among linear decision rules) because the computational burden of finding *any* function is heavy.
- The value $u_t = \psi_t(x_t)$ can be computed at time t
  - either exactly by solving a proper optimization problem, which raises issues of dynamic consistency,
  - or approximately (for example, by assuming that controls from time t on are open-loop).

The dynamic programming (DP) equation is the main theoretical tool to find optimal strategies in stochastic control problems. Dynamic programming relies on the notion of *"state"*, a summary of history sufficient for optimization [6, 78, 53, 12, 22]. However, the *"curse of dimensionality"* is an obstacle to successfully solve large size stochastic control problems.

This is why methods have been developed to tackle the state dimension issue, such as open-loop feedback optimization [12], multi-armed bandits and decentralized index policies [39], neuro-dynamic programming [14], approximate dynamic programming [51], stochastic dual dynamic programming [43], etc.

More the scenario-based approach

- An optimal "solution" can be computed scenario by scenario, with the problem that we obtain solutions such that

$$\left(w_{t_0}^s, \dots, w_{t-1}^s\right) = \left(w_{t_0}^{s'}, \dots, w_{t-1}^{s'}\right) \text{ and } u_t^s \neq u_t^{s'} , \tag{3.17}$$

  that is, a "solution" which is not implementable...
- Optimal solutions can be computed scenario by scenario and then merged (for example, by progressive hedging) to be forced to satisfy

$$\left(w_{t_0}^s, \dots, w_{t-1}^s\right) = \left(w_{t_0}^{s'}, \dots, w_{t-1}^{s'}\right) \Rightarrow u_t^s = u_t^{s'} . \tag{3.18}$$

- Optimal solutions can be computed on a tree, turning the original stochastic problem into a deterministic one, where the linear time has become an arborescent time.

Scenario trees form the basic ingredient to capture how information varies with time [63, 44]. Building a scenario tree is an issue in itself [33, 35]. This done, solving an optimization problem now amounts to switching from linear to arborescent time. Before, the control at time t was indexed by t, whereas now it is indexed by one of the nodes of the tree at time t. In addition to the numerical difficulty of computing the aggregated criterion and the constraints, one has to handle a number of variables exploding with time (exponential growth of a tree). The solution is a closed-loop strategy restricted to the nodes of the tree. Another practical issue is to extend the solution beyond ("between") the tree nodes, to capture all the available information.

As hinted above, it is equivalent to index a solution by the nodes of a scenario tree or by the scenarios, with additional informational constraints written as (long) sequences of equalities — stating that two decisions must coincide at time t if they share the same history up to time t. This is why scenarios-based optimization methods are popular in the operations research community, because they enjoy the power of all the algorithms available in standard Mathematical Programming, and particulary those around Linear Programming which are able to handle huge among of data related by linear equalities or inequalities.

3.2.5 Constraints

The non-anticipativity constraint, discussed above, belongs to the class of *informational constraints*, that is, those which express how the controls depend upon online information. Now, we turn to *pointwise constraints* that embody how variables and outputs are restricted to vary within subsets.

We can formulate constraints under uncertainty in different ways, so that constraints may be satisfied

- in the *robust, or worst-case, sense* (for all scenarios in a given subset of scenarios),
- *almost-surely* (with probability one),
- *in the mean*,
- *in probability* (with probability at least above a given threshold).

For instance, the equality "supply equals demand" may be required to be satisfied for all scenarios (robust or worst-case case, irrealistic in practice because of limited budget), or with 99% probability, etc. Other examples of constraints include maintaining the state of an operating device between security limits, either always (robust) or accepting limited violations (in probability).

3.2.6 Criterion

After having considered the feasibility issue with constraints, we finally discuss the *criterion* for optimality.

The main criterion to be optimized is generally the intertemporal sum of costs (to be minimized), or the intertemporal sum of profits (to be maximized). For instance, the *unit commitment problem* consists in planning the production of electric power generating units, meeting the demand, in order to minimize the operations costs. This constraint that the supplied energy should equal (or surpass) the load demand is typical of electricity networks. Other constraints are minimal and maximal generation levels for each unit, minimum up or down time, etc. Regulatory constraints restrict possible strategies to those satisfying environmental or safety requirements.

How the criterion is aggregated with respect to scenarios leads to various criteria under uncertainty: *expected, worst-case, multi-prior, Hurwicz, risk measures, etc.* In energy management under uncertainty, intertemporal costs are generally minimized in the mean (mathematical expectation under a single probability distribution). However, risk measures are progressively introduced [44, 63].

3.3 Framing Stochastic Optimization Problems

In §3.2, we have outlined the ingredients to formalize stochastic optimization problems, while keeping the mathematical material to a modest level. Now, we turn to modelling and we sketch how to assemble these ingredients to cook up manifold mathematical formulations.

3.3.1 Framing of a static stochastic optimization problem

In most problems, uncertainties abound. In stochastic optimization, these uncertainties are modeled by random variables[4] or stochastic processes, together with their joint probability distribution.[5] Selecting possible classes of probabilities, reflecting in particular dependencies between random variables, is a modelling issue. Specifying the parameters of the law is a statistical problem that has also to be dealt with, although it is not a part of the optimization problem itself.

With uncertainties, the cost itself becomes a random variable. As one cannot easily rank two random variables (when is one random cost "better" than another?), one usually averages out and aggregates the random cost to produce a single number. The most used *random aggregator* is the mean, or mathematical expectation. In some cases (financial problems), the expectation is taken with respect to another probability (namely the risk-neutral probability) than the original one, or alternative random aggregators, representing alternative risk preferences, can be used.

[4] We use *random variable* as a generic term that includes random vectors and stochastic processes. Throughout this manuscript, we write random variables in bold. We consistently use the notation $\boldsymbol{W}$ for the noises, i.e. the exogenous random variables.

[5] In a connex area known as *robust optimization* (see [10, 8]), uncertainties are modeled as sets of values that the uncertain parameters can take, and optimization is performed with respect to the worst possible case.

The traditional stochastic optimization problem is formulated as

$$\min_{\boldsymbol{U}\in\mathcal{U}^{ad}\subset\mathcal{U}} \quad \mathbb{E}\big[J(\boldsymbol{U},\boldsymbol{W})\big] \tag{3.19a}$$

$$s.t. \qquad \boldsymbol{U} \preceq \mathcal{B} \tag{3.19b}$$

where

- $(\Omega, \mathcal{F}, \mathbb{P})$ is a probability space, and $\mathbb{E}$ is the mathematical expectation;
- $\mathcal{U}$ is the space of all random variables $\boldsymbol{U} : \Omega \to \mathbb{U}$, where $\mathbb{U}$ is a measurable space;
- $\boldsymbol{W} : \Omega \to \mathbb{W}$ is a random variable that represents *exogenous noise*, where $\mathbb{W}$ is a measurable space;
- $J : \mathbb{U} \times \mathbb{W} \to \mathbb{R}$ is the objective function, defined on the product set of controls and uncertainties;
- $\mathcal{B} \subset \mathcal{F}$ is a sigma-algebra, and the notation $\boldsymbol{U} \preceq \mathcal{B}$ stands for "$\boldsymbol{U}$ is a random variable measurable with respect to the sigma-algebra $\mathcal{B}$", namely

  $$\boldsymbol{U} \preceq \mathcal{B} \quad \Longleftrightarrow \quad \sigma(\boldsymbol{U}) \subset \mathcal{B}\,, \tag{3.20}$$

  and captures *measurability* or *information* constraints; intuitively, the sigma-algebra $\mathcal{B}$ represents the information available to the decision-maker when choosing the control $\boldsymbol{U}$;
- $\mathcal{U}^{ad}$ is a subset of $\mathcal{U}$ that represents all remaining *constraints* like *set membership* constraints (say, inequality or equality constraints), *risk constraints*, etc.

We wish to highlight the specificities of *stochastic* optimization w.r.t. *deterministic* optimization. In this perspective, we focus on the information constraints, and we lay out different ways to represent them mathematically. Instead of the "algebraic formulation" (3.20), we can use an almost-sure equality

$$\boldsymbol{U} - \mathbb{E}\big[\boldsymbol{U} \;\big|\; \mathcal{B}\big] = 0\,, \;\; \mathbb{P}-a.s.\,, \tag{3.21}$$

which expresses that the random variable $\boldsymbol{U}$ is equal to its conditional expectation w.r.t. $\mathcal{B}$, and is thus $\mathcal{B}$-measurable.

When the sigma-algebra $\mathcal{B}$ is generated by a random variable $\boldsymbol{X} : \Omega \to \mathbb{X}$, that is, when $\mathcal{B} = \sigma(\boldsymbol{X})$, and when $\mathbb{U}$ is a separable complete metric space, a result due to J. Doob (see [23, Chapter 1, p. 18]) states that $\boldsymbol{U} \preceq \boldsymbol{X}$ is equivalent to the existence of a measurable function $\pi : \mathbb{X} \to \mathbb{U}$ such that $\boldsymbol{U} = \pi(\boldsymbol{X})$. Thus, we obtain a "functional formulation" of an information constraint:

$$\boldsymbol{U} \preceq \sigma(\boldsymbol{X}) \quad \Longleftrightarrow \quad \exists \pi : \mathbb{X} \to \mathbb{U} \text{ measurable, such that } \boldsymbol{U} = \pi(\boldsymbol{X})\,. \tag{3.22}$$

Note that this this functional formulation captures the intuitive idea associated with an information constraint: if two different elements ω and ω' of Ω cannot be distinguished by

the only observation of the information variable $\boldsymbol{X}$, then the decision-maker has to use the same control for these two elements:

$$\boldsymbol{X}(\omega) = \boldsymbol{X}(\omega') \quad \Rightarrow \quad \boldsymbol{U}(\omega) = \boldsymbol{U}(\omega') \ .$$

We distinguish two notions of *solution*, depending on the sigma-algebra $\mathcal{B}$ in (3.19b).

Open-Loop. An *open-loop solution* is $\boldsymbol{U} \preceq \{\emptyset, \Omega\}$, that is, a constant random variable. Then, the random variable is represented by its unique value.

Closed-Loop. By contrast, a *closed-loop solution* may depend on the uncertainty: $\boldsymbol{U} \preceq \mathcal{B}$, where $\{\emptyset, \Omega\} \subsetneq \mathcal{B} \subset \mathcal{F}$.

3.3.2 Multistage stochastic optimization problem

By contrast with static stochastic problems, a multistage stochastic problem introduces *stages* — labeled with integers $t = 0, \dots, T-1$, with *horizon* $T \geq 2$ — and several measurability constraints instead of only one in (3.19b). The general multistage stochastic optimization problem reads

$$\min_{(\boldsymbol{U}_0, \dots, \boldsymbol{U}_{T-1}) \in \mathcal{U}^{ad} \subset \mathcal{U}} \quad \mathbb{E}\big[J(\boldsymbol{U}_0, \cdots, \boldsymbol{U}_{T-1}, \boldsymbol{W})\big] \tag{3.23a}$$

$$s.t. \quad \boldsymbol{U}_t \preceq \mathcal{B}_t \ , \qquad \forall t \in [\![0, T-1]\!] \ , \tag{3.23b}$$

where

- $(\Omega, \mathcal{F}, \mathbb{P})$ is a probability space, and $\mathbb{E}$ is the mathematical expectation;
- $\mathcal{U}$ is the space of all random variables $(\boldsymbol{U}_0, \cdots, \boldsymbol{U}_{T-1}) : \Omega \to \mathbb{U}_0 \times \cdots \times \mathbb{U}_{T-1}$, where all $\mathbb{U}_t$ are measurable spaces;
- $\boldsymbol{W} : \Omega \to \mathbb{W}$ is a random variable that represents *exogenous noise*, where $\mathbb{W}$ is a measurable space;
- $J : \mathbb{U}_0 \times \cdots \times \mathbb{U}_{T-1} \times \mathbb{W} \to \mathbb{R}$ is the objective function;
- $\mathcal{B}_t \subset \mathcal{F}$ is a sigma-algebra, for $t \in [\![0, T-1]\!]$, and the condition $\boldsymbol{U}_t \preceq \mathcal{B}_t$ captures *measurability* or *information* constraints at stage t;
- $\mathcal{U}^{ad}$ is a subset of $\mathcal{U}$ that represents all remaining *constraints*, including ones that connect different stages.

Now, we wish to highlight the specificities of multistage *stochastic* optimization w.r.t. multistage *deterministic* optimization in a setting where information flows sequentially. We distinguish two notions of *solution*, depending on the sigma-algebras $\mathcal{B}_0, \dots, \mathcal{B}_{T-1}$.

Open-Loop. An *open-loop solution* is $(\boldsymbol{U}_0, \dots, \boldsymbol{U}_{T-1})$ such that $\boldsymbol{U}_t \preceq \{\emptyset, \Omega\}$ for all $t \in [\![0, T-1]\!]$, that is, $\boldsymbol{U}_t$ is a constant random variable.

Closed-Loop. By contrast, a *closed-loop solution* may depend on the uncertainty when $\{\emptyset, \Omega\} \subsetneq \mathcal{B}_t \subset \mathcal{F}$ for at least one $t \in [\![0, T-1]\!]$.

The case of *information accumulation* — also called *perfect memory* — is grasped with the inclusions

$$\mathcal{B}_0 \subset \cdots \subset \mathcal{B}_{T-1} \ . \tag{3.24}$$

Note that the inclusion $\mathcal{B}_t \subset \mathcal{B}_{t+1}$ means that the information $\mathcal{B}_t$ available at time t is included in the information $\mathcal{B}_{t+1}$ available at time $t+1$, hence the memory effect.

Until now, we did not require that the exogenous noise $\boldsymbol{W}$ be a sequence $\{\boldsymbol{W}_0, \dots, \boldsymbol{W}_{T-1}\}$. But, when $\boldsymbol{W}$ is a random process, we can capture the property of *non-anticipativity* by

$$\forall t \in [\![0, T-1]\!] \ , \qquad \mathcal{B}_t \subset \sigma(\boldsymbol{W}_0, \dots, \boldsymbol{W}_t) \ . \tag{3.25}$$

In the sequel, we denote by $\mathcal{F}_t$ the sigma-algebra

$$\mathcal{F}_t = \sigma(\boldsymbol{W}_0, \dots, \boldsymbol{W}_t) \ ,$$

generated by the noises up to time t, which represents the maximal information available at time t in the non-anticipative case.

The formalism (3.23b) covers the case where $\mathcal{B}_t \subset \mathcal{F}$ does not depend on past controls $\boldsymbol{U}_0, \dots, \boldsymbol{U}_{t-1}$ (like $\mathcal{B}_t = \sigma(\boldsymbol{W}_0, \dots, \boldsymbol{W}_t)$), and the case where $\mathcal{B}_t \subset \mathcal{F}$ indeed depends on past controls $\boldsymbol{U}_0, \dots, \boldsymbol{U}_{t-1}$ (like $\mathcal{B}_t = \sigma(\boldsymbol{U}_0, \dots, \boldsymbol{U}_{t-1})$).

The two most important multistage stochastic optimization theories can be distinguished according to how they handle the information constraints (3.23b):

- in the Stochastic Programming framework, the information is generally encoded in a tree, and the sigma-algebra $\mathcal{B}_t$ corresponds to the set of nodes at stage t;
- in the Stochastic Optimal Control framework, the sigma-algebra $\mathcal{B}_t$ is $\sigma(\boldsymbol{X}_t)$ generated by an information state $\boldsymbol{X}_t$, produced by a controlled dynamics.

Both theories incorporate a non-anticipativity property, as well as information accumulation (under the Markovian setup in Stochastic Optimal Control). We now present Stochastic Programming and Stochastic Optimal Control with a focus on the information constraints (3.23b).

Stochastic Programming (SP)

In Stochastic Programming, the probability space $(\Omega, \mathcal{F}, \mathbb{P})$ is called *scenario space*, where *scenarios* stand for sequences of *uncertainties*. The sequential structure of information arrival about uncertainty is represented either by a subset of a product space or by a so-called *scenario tree* (see Figure 3.1).

For the sake of simplicity, in this manuscript we only consider Stochastic Programming for finite scenario spaces. For a set of scenario Ω we suppose given

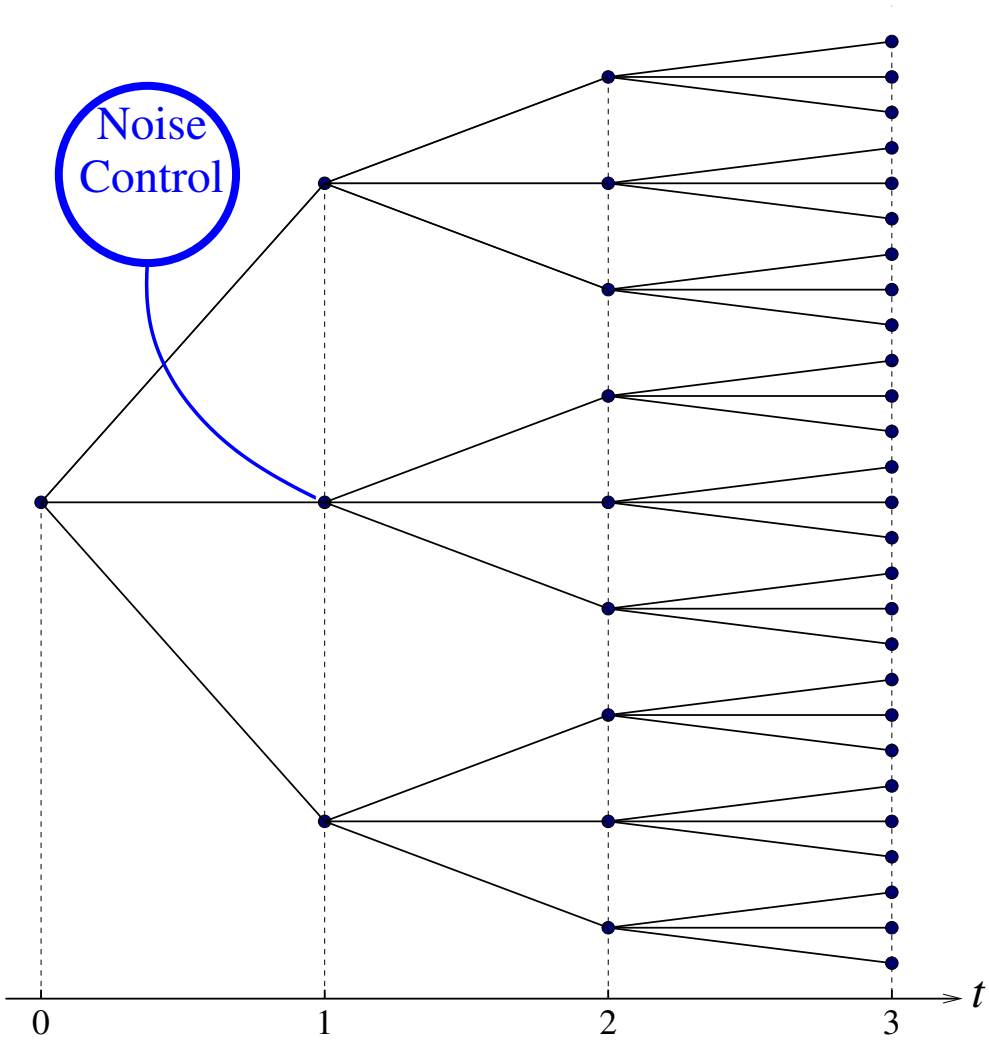


Figure 3.1: A scenario tree

- a probability $\mathbb{P}$ on Ω;
- control sets $\mathbb{U}_0, \ldots, \mathbb{U}_{T-1}$;
- an uncertainty set $\mathbb{W}$ and a mapping $\boldsymbol{W} : \Omega \to \mathbb{W}$ that represents *exogenous noises*;
- an objective function $J : \mathbb{U}_0 \times \cdots \times \mathbb{U}_{T-1} \times \mathbb{W} \to \mathbb{R}$.

Stochastic Programming with scenario space. In Stochastic Programming, the finite probability space $(\Omega, \mathcal{F}, \mathbb{P})$ can be represented as a subset of a product space

$$\Omega \subset \Omega_0 \times \cdots \times \Omega_{T-1} \ , \tag{3.26a}$$

where the set Ω_t supports the uncertainties at step t, so that a scenario is denoted by

$$\omega = (\omega_0, \ldots, \omega_{T-1}) = \{\omega_s\}_{s=0}^{T-1} \ . \tag{3.26b}$$

A possible solution is a family of controls $u_t(\omega) \in \mathbb{U}_t$ doubly indexed by step t and uncertainty ω. The non-anticipativity constraint (Constraint (3.23b) where $\mathcal{B}_t = \mathcal{F}_t$) is captured by the requirement that, for all $t \in [\![0, T-1]\!]$,

$$\forall (\omega, \omega') \in \Omega^2 \ , \qquad \{\omega_s\}_{s=0}^t = \{\omega'_s\}_{s=0}^t \quad \Longrightarrow \quad u_t(\omega) = u_t(\omega') \ . \tag{3.27}$$

The general stochastic programming problem reads

$$\begin{aligned} \min_{\{\{u_t(\omega)\}_{\omega\in\Omega}\}_{t=0}^{T-1}} \quad & \sum_{\omega\in\Omega} \mathbb{P}\big(\{\omega\}\big) J\Big(\{u_t(\omega)\}_{t=0}^{T-1}, \boldsymbol{W}(\omega)\Big) \ . \\ s.t. \quad & \text{constraint (3.27)} \end{aligned} \tag{3.28}$$

We develop in Table 3.1 the correspondence between the framing of Stochastic Programming problems with scenario space and the abstract framing of §3.3.2.

	Stochastic Programming formulation	Abstract formulation
States of Nature	$\Omega \subset \Omega_0 \times \cdots \times \Omega_{T-1}$ finite set	Ω measurable space
Probability	$\{\mathbb{P}\big(\{\omega\}\big)\}_{\omega\in\Omega}$	$\mathbb{P}$
Solution	$\{\{u_t(\omega)\}_{\omega\in\Omega}\}_{t=0}^{T-1}$ $\forall \omega \in \Omega \ , \ \forall t \in [\![0, T-1]\!]$, $\{\omega_s\}_{s=0}^t = \{\omega'_s\}_{s=0}^t \Rightarrow u_t(\omega) = u_t(\omega')$	$\{\boldsymbol{U}_t\}_{t=0}^{T-1}$ $\forall t \in [\![0, T-1]\!]$, $\boldsymbol{U}_t \preceq \mathcal{F}_t$

Table 3.1: Correspondence between Stochastic Programming with scenario space framing and abstract framing

From nested partitions to scenario tree. The stochastic programming community often presents optimization problems formulated on a scenario tree. We give a formal definition of a scenario tree attached to a finite set equipped with a sequence of nested partitions.

Definition 3.1 *Consider a set Ω (not necessarily included in a product space), equipped with a sequence $\{\mathcal{N}_t\}_{t=0}^{T-1}$ of partitions such that $\mathcal{N}_{t+1}$ is a refinement of $\mathcal{N}_t$ (i.e. any element of $\mathcal{N}_{t+1}$ is contained in an element of $\mathcal{N}_t$). We call $\mathcal{T} = \{\mathcal{N}_t\}_{t=0}^{T-1}$ a* scenario forest*. A scenario forest is called a* scenario tree *in the case where $\mathcal{N}_0 = \{\Omega\}$ and $\mathcal{N}_{T-1} = \big\{\{\omega\} \mid \omega \in \Omega\big\}$.*

Hence, on a scenario tree, a scenario $\omega \in \Omega$ is associated with a leaf $\{\omega\} \in \mathcal{N}_{T-1}$ *of the tree. A* node of depth t *of the tree $\mathcal{T}$ is an element of $\mathcal{N}_t$. A node n is said to be an* ascendant *of a node m if $m \subset n$, and we denote by $a(m)$ the set of ascendant nodes of m. Conversely, m is a* descendent *of n and, with a node $n \in \mathcal{N}_t$ we associate the set $c(n)$ of its* children nodes*, consisting of the nodes $m \in \mathcal{N}_{t+1}$ that are descendent of n. The* genealogy *of a node is the collection of all its ascendants.*

We also define the functions $n_t : \Omega \to \mathcal{N}_t$ satisfying $\omega \in n_t(\omega)$.

Now, we can proceed to detail links between stochastic programming formulations with and without scenario tree.

Note that, if the finite set Ω carries a probability $\mathbb{P}$, we deduce the probability of each node $n \in \mathcal{T}$ thanks to the construction in Definition 3.1:

$$\mathbb{P}(n) = \sum_{\omega \in n} \mathbb{P}(\{\omega\}) \ .$$

This formula may be expressed in a recursive way by using the sets of children nodes, namely

$$\mathbb{P}(n) = \sum_{m \in c(n)} \mathbb{P}(m) \ .$$

In the case where $\Omega \subset \Omega_0 \times \cdots \times \Omega_{T-1}$, we can construct a tree thanks to the existence of a natural sequence $\{\mathcal{N}_t\}_{t=0}^{T-1}$ of nested partitions. Indeed, for all $t \in [\![0, T-1]\!]$, we define $\mathcal{N}_t$ as the equivalence classes for the relation (of common past)

$$\omega \sim_t \omega' \iff \{\omega_s\}_{s=0}^{t} = \{\omega'_s\}_{s=0}^{t} \ , \tag{3.29}$$

where $\omega = \{\omega_s\}_{s=0}^{T-1}$ and $\omega' = \{\omega'_s\}_{s=0}^{T-1}$ are any elements of Ω. Conversely, we easily construct a product set of uncertainties from a tree, and identify the tree with a subset (see Figure 3.2).

Stochastic Programming with scenario tree. In stochastic programming formulations of optimization problems with a scenario tree, a possible solution is a family of controls $\Big\{\big\{u_{n_t}\big\}_{n_t \in \mathcal{N}_t}\Big\}_{t=0}^{T-1}$ indexed by the nodes of the tree, where, for any time t, and any node $n_t \in \mathcal{N}_t$, $u_{n_t} \in \mathbb{U}_t$. In this way, the information constraints (3.23b) are automatically captured in the very indexing of a possible solution by the nodes n_t of the tree: at step t, a solution can only depend on past uncertainties $\omega_0, \ldots, \omega_t$.

With this notion of solution, the general stochastic programming problem reads

$$\min_{\{u_n\}_{n \in \mathcal{T}}} \sum_{\omega \in \Omega} \mathbb{P}\big(\{\omega\}\big) J\Big(\big\{u_n\big\}_{n \in a(\{\omega\})}, W\big(\omega\big)\Big) \ . \tag{3.30}$$

We shall be brought to consider a specific class of problems, additive in time and with a dynamics:

$$\min_{\big\{\{u_{n_t}\}_{n_t \in \mathcal{N}_t}\big\}_{t=0}^{T-1}} \quad \sum_{t=0}^{T-1} \sum_{n_t \in \mathcal{N}_t} \sum_{m \in c(n_t)} \mathbb{P}\big(\{m\}\big) \ L_t\Big(x_{n_t}, u_{n_t}, w_m\Big) \tag{3.31a}$$

$$s.t. \quad x_m = F_t\big(x_{n_t}, u_{n_t}, w_m\big) \ , \qquad \forall m \in c(n_t) \ , \ \forall n_t \in \mathcal{N}_t \ , \ \forall t \ . \tag{3.31b}$$

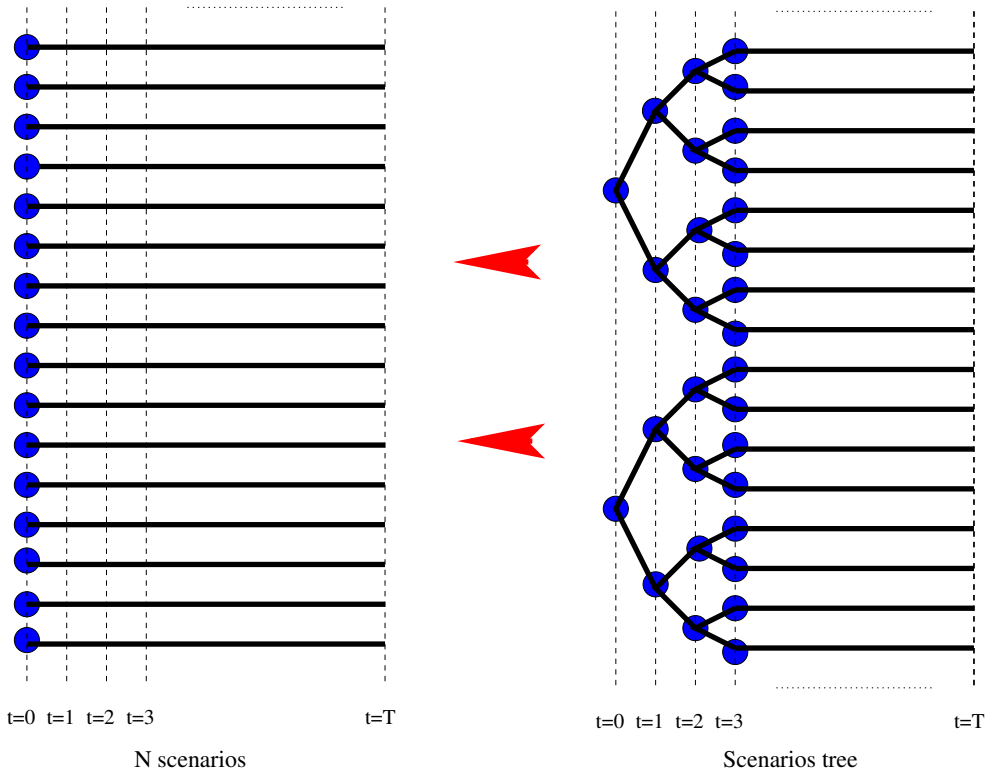


Figure 3.2: From scenario tree to set of scenarios

We develop in Table 3.2 the correspondence between the framing of Stochastic Programming problems with scenario tree and the abstract framing of §3.3.2.

Stochastic Optimal Control (SOC)

In Stochastic Optimal Control, the information constraints (3.23b) are materialized by means of a so-called *state*. The framing comprises a *Stochastic Dynamic System* (SDS) consisting of

- a sequence $\{\mathbb{X}_t\}_0^T$ of sets of *states*;
- a sequence $\{\mathbb{U}_t\}_0^{T-1}$ of sets of *controls*;
- a sequence $\{\mathbb{W}_t\}_0^{T-1}$ of sets of *uncertainties*,;
- a sequence $\{F_t\}_0^{T-1}$ of *functions*, where $F_t : \mathbb{X}_t \times \mathbb{U}_t \times \mathbb{W}_t \to \mathbb{X}_{t+1}$, play the role of *dynamics* at time t;
- a probability space $(\Omega, \mathcal{F}, \mathbb{P})$;
- exogenous noises $\{\boldsymbol{W}_t\}_{t=0}^{T-1}$, where each $\boldsymbol{W}_t$ takes values in $\mathbb{W}_t$;

	Stochastic Programming formulation	Abstract formulation
States of Nature	$\mathcal{T}$ / $\mathcal{N}_{T-1}$ tree (forest) / leaves	Ω measurable space
Information	$\mathcal{N}_t$ nodes at time t	$\mathcal{F}_t$ sigma-algebra
Probability	$\{\mathbb{P}(\{n\})\}_{n\in\mathcal{N}_{T-1}}$	$\mathbb{P}$
Solution	$\big\{\{u_n\}_{n\in\mathcal{N}_t}\big\}_{t=0}^{T-1}$	$\{\boldsymbol{U}_t\}_{t=0}^{T-1}$ $\boldsymbol{U}_t \preceq \mathcal{F}_t$, $\forall t \in [\![0, T-1]\!]$

Table 3.2: Correspondence between Stochastic Programming with scenario tree framing and abstract framing

- an objective function $J : \mathbb{X}_0 \times \cdots \times \mathbb{X}_T \times \mathbb{U}_0 \times \cdots \times \mathbb{U}_{T-1} \times \mathbb{W}_0 \times \cdots \times \mathbb{W}_{T-1} \to \mathbb{R}$.

The sigma-algebras (3.3.2) form the filtration $\{\mathcal{F}_t\}_{t=0}^T$ of past noises. For an $\Big(\{\mathcal{F}_t\}_{t=0}^T\Big)$-adapted sequence $\{\boldsymbol{U}_t\}_{t=0}^{T-1}$ of controls — that is, random variables $\boldsymbol{U}_t$ with value in $\mathbb{U}_t$, and such that $\boldsymbol{U}_t \preceq \mathcal{F}_t$ — and an initial state $\overline{x}_0 \in \mathbb{X}_0$, we obtain a sequence $\{\boldsymbol{X}_t\}_{t=0}^T$ of states as follows:

$$\forall t \in [\![0, T-1]\!] \,, \qquad \boldsymbol{X}_{t+1} = F_t\big(\boldsymbol{X}_t, \boldsymbol{U}_t, \boldsymbol{W}_t\big) \,. \tag{3.32}$$

We observe that, for any time $t \in [\![1, T]\!]$, $\boldsymbol{X}_t$ is measurable w.r.t. $\mathcal{F}_{t-1} \subset \mathcal{F}_t$ by construction.

We denote

$$\boldsymbol{X} = \{\boldsymbol{X}_t\}_{t=0}^T \,, \;\; \boldsymbol{U} = \{\boldsymbol{U}_t\}_{t=0}^{T-1} \,, \;\; \boldsymbol{W} = \{\boldsymbol{W}_t\}_{t=0}^{T-1} \,. \tag{3.33}$$

The general stochastic optimal control problem reads

$$\min_{\boldsymbol{X},\boldsymbol{U}} \quad \mathbb{E}\Big[J(\boldsymbol{X}, \boldsymbol{U}, \boldsymbol{W})\Big] \tag{3.34a}$$

$$s.t. \quad \boldsymbol{X}_{t+1} = F_t\big(\boldsymbol{X}_t, \boldsymbol{U}_t, \boldsymbol{W}_t\big) \qquad \forall t \in [\![0, T-1]\!] \,, \tag{3.34b}$$

$$\boldsymbol{U}_t \preceq \mathcal{B}_t \qquad \forall t \in [\![0, T-1]\!] \,, \tag{3.34c}$$

where $\mathcal{B}_t \subset \mathcal{F}_t$ is a sigma-algebra, for $t \in [\![0, T-1]\!]$, and the conditions $\boldsymbol{U}_t \preceq \mathcal{B}_t$ captures *measurability* or *information* constraints at stage t.

Here again, we wish to highlight the specificities of multistage *stochastic* optimization w.r.t. multistage *deterministic* optimization, in a setting where information flows sequentially. Since $\mathcal{B}_t \subset \mathcal{F}_t$, the condition (3.34c) implies that the control $\boldsymbol{U}_t$ is chosen knowing only the past noises $\boldsymbol{W}_0, \dots, \boldsymbol{W}_t$. This is the so-called *nonanticipativity* constraint: $\boldsymbol{U}_t$ is measurable with respect to $\mathcal{F}_t$.

We distinguish several classes of information structures, depending on $\mathcal{B}_t$ in the condition (3.34c), hence several notions of *solution*.

Open-Loop. An *open-loop solution* is one where the condition (3.34c) reads $\boldsymbol{U}_t \preceq \{\emptyset, \Omega\}$, for all $t \in [\![0, T-1]\!]$. In other words, $\mathcal{B}_t = \{\emptyset, \Omega\}$, for all $t \in [\![0, T-1]\!]$.

Closed-Loop. A solution satisfying the condition (3.34c) is a *closed loop solution* as soon as $\{\emptyset, \Omega\} \subsetneq \mathcal{B}_t \subset \mathcal{F}_t$ for at least one $t \in [\![0, T-1]\!]$. In practice, information may be provided through the observation of either the state variables, or the noise variables, or a mix; the following subdivisions are often considered.

- In the *Decision-Hazard* setting, $\mathcal{B}_t = \sigma(\boldsymbol{X}_t)$ in (3.34c) so that decisions $\boldsymbol{U}_t \preceq \boldsymbol{X}_t$ are taken before knowing the uncertainty $\boldsymbol{W}_t$ at time t, and only according to the current state $\boldsymbol{X}_t$. By the Doob result (3.22), a solution can be expressed as a *state feedback* $\boldsymbol{U}_t = \pi_t(\boldsymbol{X}_t)$, where $\pi_t : \mathbb{X}_t \to \mathbb{U}_t$.
- In the *Hazard-Decision* setting, $\mathcal{B}_t = \sigma(\boldsymbol{X}_t, \boldsymbol{W}_t)$ in (3.34c) so that decisions $\boldsymbol{U}_t \preceq \sigma(\boldsymbol{X}_t, \boldsymbol{W}_t)$ are taken after knowing the uncertainty at time t, according to the current state $\boldsymbol{X}_t$ and the current uncertainty $\boldsymbol{W}_t$. By the Doob result (3.22), a solution can be expressed as $\boldsymbol{U}_t = \pi_t(\boldsymbol{X}_t, \boldsymbol{W}_t)$, where $\pi_t : \mathbb{X}_t \times \mathbb{W}_t \to \mathbb{U}_t$.
- The largest class of closed loop solutions is of course obtained when $\mathcal{B}_t = \mathcal{F}_t$ for all $t \in [\![0, T-1]\!]$. When the exogenous noises $\{\boldsymbol{W}_t\}_{t=0}^{T-1}$ form a sequence of independent random variables, it can be shown that there is no loss of optimality in reducing the search to the class of Hazard-Decision feedback solutions, namely $\mathcal{B}_t = \sigma(\boldsymbol{X}_t, \boldsymbol{W}_t)$. When the size of the state space $\mathbb{X}_t$ does not increase with t, and neither does $\mathbb{W}_t$, this property has major consequences for numerical applications.
- A smaller class of closed loop solutions is obtained when $\mathcal{B}_t = \mathcal{F}_{t-1}$ for all $t \in [\![0, T-1]\!]$. When the exogenous noises $\{\boldsymbol{W}_t\}_{t=0}^{T-1}$ form a sequence of independent random variables, it can be shown that there is no loss of optimality in reducing the search to the class of state feedback solutions, namely $\mathcal{B}_t = \sigma(\boldsymbol{X}_t)$. When the size of the state space $\mathbb{X}_t$ does not increase with t, this property has major consequences for numerical applications.

This general form (3.34) is not common, and one generally rather considers a time additive expression for the cost function, namely,

$$\min_{\pi=\{\pi_t\}_{t=0}^{T-1}} \quad \mathbb{E}\Big[\sum_{t=0}^{T-1} L_t\big(\boldsymbol{X}_t, \boldsymbol{U}_t, \boldsymbol{W}_t\big) + K\big(\boldsymbol{X}_T\big)\Big] \tag{3.35a}$$

$$s.t. \quad \boldsymbol{X}_{t+1} = F_t\big(\boldsymbol{X}_t, \boldsymbol{U}_t, \boldsymbol{W}_t\big)\ , \qquad \forall t \in [\![0, T-1]\!]\ , \tag{3.35b}$$

$$\boldsymbol{U}_t = \pi_t(\boldsymbol{X}_t)\ ,\ \ \pi_t : \mathbb{X}_t \to \mathbb{U}_t\ , \qquad \forall t \in [\![0, T-1]\!]\ , \tag{3.35c}$$

where

- $L_t : \mathbb{X}_t \times \mathbb{U}_t \times \mathbb{W}_t \mapsto \mathbb{R}$ is the *instantaneous cost* at step t, for all $t \in [\![0, T-1]\!]$, and $K : \mathbb{X}_T \to \mathbb{R}$ is the final cost;
- the policies π_t are measurable mappings, for all $t \in [\![0, T-1]\!]$, and capture *information* constraints;

Remark 3.2 *We discuss the notion of state in §4.1.4. In the Decision-Hazard setting, $\mathcal{B}_t = \mathcal{F}_{t-1}$ in (3.34c), and the quantity $\boldsymbol{X}_t$ is said to be an* information state *when the condition $\boldsymbol{U}_t \preceq \mathcal{F}_{t-1}$ (obtained from (3.34c) with $\mathcal{B}_t = \mathcal{F}_{t-1}$) can be replaced by the condition $\boldsymbol{U}_t \preceq \boldsymbol{X}_t$, where $\boldsymbol{X}_t$ is $\mathcal{F}_{t-1}$-measurable. The same can be done in the Hazard-Decision setting.*

In Problem (3.35), the condition (3.35c) suggests that the state $\boldsymbol{X}_t$ is an information state *as the decisions are taken in function of it; we say "suggests" because this relies on the implicit assumption that there is no loss of optimality in reducing the search to the class of state feedback solutions, instead of the largest class of adapted controls. In Problem (3.34), $\boldsymbol{X}_t$ is simply the physical state (and might or might not be an information state, depending on additional assumptions).*

As just discussed, the form (3.35) is especially adapted to the case where the exogenous noises $\{\boldsymbol{W}_t\}_{t=0}^{T-1}$ form a sequence of independent random variables. We will come back to that point when we address Dynamic Programming in §4.1.4.

Connection between SP and SOC

The SOC framing includes the SP one, at the expense of introducing a state like in Table 3.3.

Observe that the state $\boldsymbol{X}_t$ at stage t is huge, as it includes all the exogenous noises $\boldsymbol{W}$ and the past controls $\boldsymbol{U}_0, \dots, \boldsymbol{U}_{t-1}$. Observe also the not common fact that the state $\boldsymbol{X}_t$ at stage t includes *all* the noises $\boldsymbol{W} = \big\{\boldsymbol{W}_0, \dots, \boldsymbol{W}_{T-1}\big\}$, be they past, present or future! As a consequence, the state $\boldsymbol{X}_t$ is *not* $\mathcal{F}_t$-measurable, hence is not observable by the decision-maker at stage t and cannot be the input of any implementable feedback. What is more, the dimension of the state grows with the stages, as reflected in the dynamics that just extends the vector x_t by adding u_t to the right: the state $\boldsymbol{X}_t$ at stage t keeps track of past controls $\boldsymbol{U}_0, \dots, \boldsymbol{U}_{t-1}$ by accumulating them. This state is called the "maximal state", and

	Stochastic Programming formulation	Stochastic Optimal Control formulation
States of Nature	$\Omega \subset \Omega_0 \times \cdots \times \Omega_{T-1}$ finite set	Ω measurable space
Exogenous noise		$\mathbb{W}_t = \Omega_t$, $\boldsymbol{W}_t : \Omega \to \Omega_t$ projection
Probability	$\{\mathbb{P}(\{\omega\})\}_{\omega\in\Omega}$	$\mathbb{P}$
State		$\boldsymbol{X}_t = (\overline{x}_0, \boldsymbol{W}, \boldsymbol{U}_0, \ldots, \boldsymbol{U}_{t-1})$
Information		$\mathcal{F}_t = \sigma(\boldsymbol{W}_0, \cdots, \boldsymbol{W}_t)$
Dynamics		$F_t(x_t, u_t, w_t) = (x_t, u_t)$

Table 3.3: Turning a Stochastic Programming framing into Stochastic Optimal Control framing

it will again be discussed in §4.1.4. This is not an information state as it is not totally observable (see Remark 3.2), whereas we will see that the conditional distribution of the maximal state $\boldsymbol{X}_t$ knowing $\mathcal{F}_t$ is. In practice, depending on the specificities of the model, it may happen that smaller states can be displayed.

3.3.3 Discussion of complexity

Finally, we briefly point out in what sense multistage stochastic optimization problems are complex, as a motivation to the coming Chapter 4, dedicated to resolution methods.

To give a feeling of the complexity of a multistage stochastic optimization problem, we assume that controls take their values in a finite set of cardinal n_u. Therefore, there are

$$(n_u)^{T|\Omega|} \tag{3.36}$$

possible solutions (not all of them are admissible).

To account for non-anticipativity and restrict solutions, we suppose that the sample space Ω is a product of t copies with cardinal n_w, so that $|\Omega| = (n_w)^T$. Hence, the number of possible solutions is

$$(n_u)^{T(n_w)^T}\ , \tag{3.37}$$

and the number of non-anticipative ones is

$$(n_u)^{\sum_{s=0}^{T-1}(n_w)^s} = (n_u)^{\frac{(n_w)^T-1}{n_w-1}}\ . \tag{3.38}$$

This number is also the number of possible solutions when the set Ω is represented by the leaves of a tree of depth T, each node having n_w children, because then the number of nodes is $\sum_{s=0}^{T-1} n_w^s = \frac{(n_w)^T-1}{n_w-1}$.

In almost any practical situation, this is of course a huge number and something has to be done in order to compress it: see for example Remark 4.4 devoted to the complexity of Dynamic Programming in Chapter 4. Further discussion on the complexity of multistage stochastic optimization problems can be found in [64, 67].

3.4 Extensions to More General Risk Attitudes

A large body of the literature, not to say the largest, deals with what economists coin the *risk neutral* approach, where the random criterion is aggregated by means of the mathematical expectation. Now, issues such as environmental restrictions, safety rules, etc. force us to introduce certain types of constraints (in probability) and risk (risk measures) in the optimization framework. This is why we discuss various approaches, where the decision-maker is compelled to explicit his/her attitudes with respect to risk. Our main references are [60, 26].

3.4.1 How can we rank random variables?

In a probabilistic setting, where uncertainties are random variables, the output of a stochastic optimization problem is a random variable, the optimal random criterion. The issue is: how can we rank random variables? In practice, this question revolves around defining how to handle uncertainties in the payoff criterion and in the constraints.

Risk neutral

The risk neutral approach states that random variables are ranked by being reduced to their means. In most stochastic optimization problems formulations, one takes the mathematical expectation of the criterion and asks that pointwise constrainsts be satisfied almost surely, that is, practically, for all possible outcomes of the uncertainties (robust approach).

However, there are many other ways to handle risk: robust, worse case, risk measures, in probability, almost surely, etc.

Robust, worst case, pessimistic, maximin, minimax

In the *robust* approach, the DM considers that

- all he knows about scenarios is that they are restricted to a known given subset of the set of scenarios,
- Nature is supposed to be "malevolent", and specifically selects the worst scenario after the DM has made his decision.

In this pessimistic approach, the DM will select solutions that make the best against the *worst case*, hence the vocables of *maximin* and *minimax*.

The robust approach can be softened with plausibility weightings, that are the additive equivalents of probabilities in the stochastic approach. In [11], Pierre Bernhard coined *fear operator* the worst-case operator, widely considered in the field of robust optimization, to contrast with the *mathematical expectation operator*, widely considered in the field of stochastic optimization.

Optimistic, Hurwicz

If you follow Ambrose Bierce's tongue-in-cheek definition of future as "That period of time in which our affairs prosper, our friends are true and our happiness is assured", then you should adopt the optimistic approach. As Nature is supposed to "benevolent", instead of optimizing the worst output as in the robust approach, the *optimistic* focuses on the best output.

Optimistic and pessimistic are the two extremities of the *Hurwicz* aggregator, which is a convex combination of the worst and the best possible values of the criterion, over scenarios.

Expected utility

In the stochastic or expected approach, Nature is supposed to play stochastically under a known probability distribution. In economics, following von Neumann and Morgenstern, you assess the criterion by means of a *utility* function, before taking the expectation. In a temporal setting, economists usually consider the *discounted expected utility*.

Ambiguity, multi-prior

The ambiguity or multi-prior approach combines robust and expected standpoints. The DM considers that

- the set of scenarios is equipped with different probabilities — termed as *beliefs* or *priors* — that belong to a known given set $\mathcal{P}$,
- under any probability $\mathbb{P} \in \mathcal{P}$, Nature is supposed to play stochastically.

The *ambiguity* about which of the different probabilities in $\mathcal{P}$ is the proper probability compels the DM to

1. first, evaluate the expected criterion for all possible probability $\mathbb{P} \in \mathcal{P}$,
2. second, take the worst among all expectations, as $\mathbb{P}$ varies in $\mathcal{P}$.

Convex risk measures

Convex risk measures cover a wide range of risk criteria: *risk-neutral expectation*, *worst case*, *ambiguity*, *tail, or average value-at-risk*. Indeed, they can be interpreted as an extension of ambiguity where the DM considers that

- the set of scenarios is equipped with different probabilities — termed as *beliefs* or *priors* — that belong to a known given set $\mathcal{P}$,
- probabilities differ between themselves according to an "additive weight" $\Theta(\mathbb{P})$,
- under any probability $\mathbb{P} \in \mathcal{P}$, Nature is supposed to play stochastically.

The *ambiguity* about which of the different probabilities in $\mathcal{P}$ is the proper probability compels the DM to

1. first, evaluate the expected criterion for all possible probability $\mathbb{P} \in \mathcal{P}$,
2. second, add the "weight" $\Theta(\mathbb{P})$ to each expectation,
3. third, take the worst among all "additive weighted" expectations, as $\mathbb{P}$ varies in $\mathcal{P}$.

Non-convex risk measures also play an important role in practice, as the *value-at-risk*. Risk measures can be used to aggregate the criterion, as well as the constraints.

3.4.2 Markowitz's mean-variance

Markowitz's approach incorporates risk in the risk neutral approach, not by the utility assessment of a random prospect, but by requiring the constraint that variance be bounded above.

3.4.3 Probability or chance constraints

A probability constraint is one that forces a stochastic process to remain within a given region with a minimal probability. Chance constrained optimization problems were introduced by [18] with an individual chance constraint and by [40] with a joint chance constraint. Such problems raise theoretical and numerical difficulties: indeed, it is mathematically difficult to guarantee the connectedness, the convexity or the closedness of the feasible set induced by the chance constraint, whereas these properties play key roles in optimization.

Chapter 4

Displaying Stochastic Optimization Resolution Methods

After having laid out, in Chapter 3, the framing of optimization problems, we are now ready to showcase resolution methods. As optimization is challenged by the complexity of the emerging power systems — displaying large size, dynamical aspects, and uncertainties — we feel that decomposition approaches may prove particularly adapted. In §4.1, we present, in an unified framework, the main approaches to decompose multistage stochastic optimization problems for numerical resolution. This framework covers both *Stochastic Programming* and *Stochastic Dynamic Programming*, the two most well-known methods in stochastic optimization. This done, we go in more detail and outline more specific ones like *Progressive Hedging* in §4.2, *Stochastic Dual Dynamic Programming (SDDP)* in §4.3, *Approximate Dynamic Programming (ADP)* in §4.4. As we cannot cover all methods, we invite international colleagues to complete the picture and to share experiences, works and practices in Chapter 5.

4.1 Resolution by Decomposition Methods in Multistage Stochastic Optimization

We present, in an unified framework, the main approaches to decompose multistage stochastic optimization problems for numerical resolution.

To fix ideas and simplify the exposition, we present a setting where all variables are parametrized by discrete indexes. For this purpose, suppose given a finite horizon T (so that time $t \in [\![0, T]\!]$), a finite probability space $\left(\Omega, \mathcal{F}, \mathbb{P}\right)$, endowed with a filtration $\{\mathcal{F}_t\}_0^{T-1}$, a finite number N of units (space). We consider the multistage stochastic optimization

problem

$$\min_{\boldsymbol{X},\boldsymbol{U}} \quad \sum_{\omega\in\Omega}\sum_{i=1}^{N}\sum_{t=0}^{T-1} \mathbb{P}\big(\{\omega\}\big)\; L_t^i\Big(\boldsymbol{X}_t^i(\omega), \boldsymbol{U}_t^i(\omega), \boldsymbol{W}_t(\omega)\Big) \tag{4.1a}$$

$$s.t. \quad \boldsymbol{X}_{t+1}^i(\omega) = F_t^i\big(\boldsymbol{X}_t^i(\omega), \boldsymbol{U}_t^i(\omega), \boldsymbol{W}_t(\omega)\big) \qquad \forall t\ ,\ \forall i\ ,\ \forall\omega \tag{4.1b}$$

$$\sum_{i=1}^{N} \theta_t^i\big(\boldsymbol{X}_t^i(\omega), \boldsymbol{U}_t^i(\omega)\big) = 0 \qquad \forall t,\quad \forall\omega \tag{4.1c}$$

$$\boldsymbol{U}_t^i \preceq \mathcal{F}_t \qquad \forall t\ ,\ \forall i\ , \tag{4.1d}$$

where ω is a scenario of uncertainties given by $\omega = \big\{\omega_t\big\}_{t=0}^{T-1}$. The constraint (4.1b) represents the dynamics of each subsystem, the constraint (4.1c) represents the coupling constraint between the subsystems (also called units), and the constraint (4.1d) is the non-anticipativity constraint. Constraints function θ_t^i are assumed to have image in $\mathbb{R}^{n_c}$.

As we have seen in §3.3.2 that the SOC framing includes the SP one, the above setting applies both to SP and SOC problems.

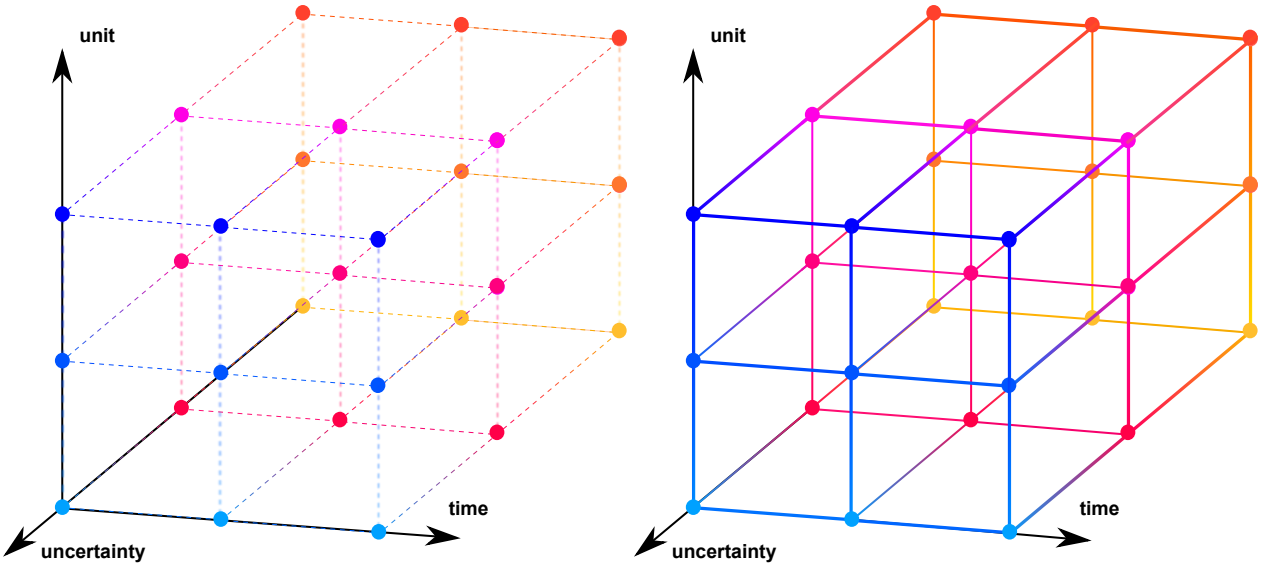
In Problem (4.1), we have local costs — depending on step t, uncertainty ω and unit i — and we minimize their sum over time, uncertainty and space. Without constraints (4.1b)-(4.1d), Problem (4.1) (illustrated in Figure 4.1a) consists in minimizing a sum of independent costs. Hence, the minimum of the sum is the sum of the minimums, and the problem is decomposed. However, the local costs are linked (illustrated in Figure 4.1b)

- in time through the dynamic of the system (e.g. Equation (4.1b));
- in unit through the coupling constraints (e.g. Equation (4.1c));
- and in scenario (uncertainty) through the nonanticipativity constraint (e.g. Equation (4.1d)).

We now lay out different ways to divide the original complex problem into easier to solve subproblems. We propose three angles to decompose the original problem: decomposition in time (step), decomposition in scenario (uncertainty) and decomposition in space (unit), as illustrated in Figure 4.2.

Moreover, we distinguish two types of decomposition.

- In *chained decomposition*, like Dynamic Programming (see [6, 12]), the original problem is solved by means of successive smaller subproblems, solved one after the other (in Dynamic Programming, each subproblem is solved only once). Chained decomposition relies on a specific structure of the coupling constraint, like the flow of time.
- In *parallel decomposition*, like Progressive Hedging (see [56, 75]), the original problem is solved by means of parallel smaller subproblems, coordinated and updated by a master algorithm. These subproblems can be obtained by dualizing the constraint, and have to be solved several times before obtaining an optimal solution to the global problem.



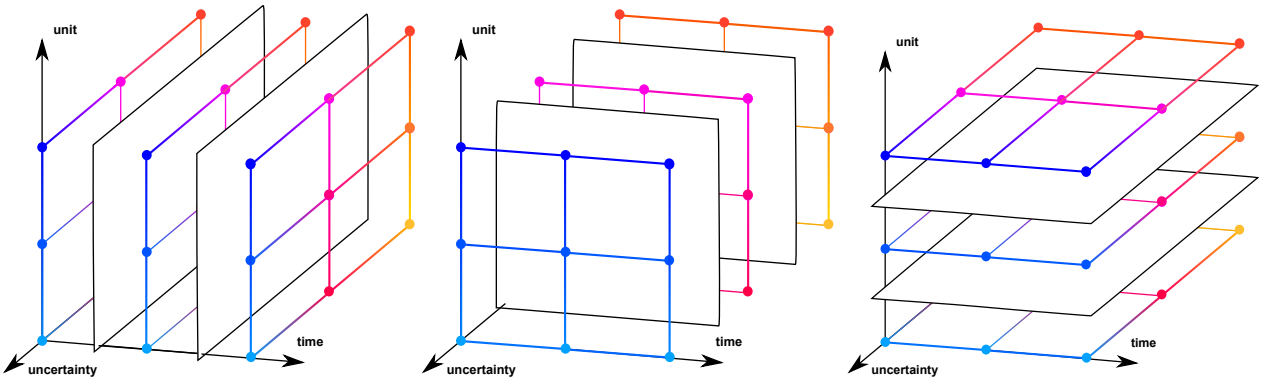
(a) Local costs (b) Local costs linked

Figure 4.1: Representation of the local costs depending on time, uncertainty (scenario) and space (unit) and the links induced by the constraints

4.1.1 Duality and parallel decomposition

Before presenting the different decompositions approaches, we now illustrate how the duality theory leads to decomposition schemes. We present here, in a simple setting, the most usual, known as *price decomposition scheme*. For clarity, the units coupling functions θ^i in (4.1c) are assumed, here, to be real valued.

This price decomposition scheme can be intuitively understood as follows. We consider a problem where a team of N units — each of them producing a quantity $\theta_i(u_i)$ function of the local control u_i — has to meet a given demand. Each unit incurs a local cost $L_i(u_i)$, and the problem consists in minimizing the sum of the local costs. The decomposition is obtained by replacing the "production equal demand" equality by a price mechanism. To achieve a proper price, we suppose that a coordinator can impose costs to all units iteratively. At iteration k, the coordinator sets a price $p^{(k)} = -\lambda^{(k)}$ for the output of each unit $\theta_i(u_i)$. Each unit then minimizes the sum of its local production cost $L_i(u_i)$ minus the cash flow produced by the output $p^{(k)}\theta_i(u_i)$, and obtains a solution $u_i^{(k)}$. Then, the coordinator collects the production of all units, makes the sum and compares the result to the demand. If the total production is not enough, he increases the price of the output; if the total production exceeds the demand, he decreases the price.



(a) Time decomposition (b) Uncertainty decomposition (c) Space decomposition

Figure 4.2: Decomposition according to time, uncertainty (scenario) or space (unit). Each plane carries a problem with coupling in only two dimensions.

More precisely, we consider the following problem:

$$\min_{\{u_i\}_{i=1}^N} \quad \sum_{i=1}^N L_i(u_i) \tag{4.2a}$$

$$s.t. \quad u_i \in U_i^{ad} \ , \qquad \forall i \in [\![1, N]\!] \ , \tag{4.2b}$$

$$\sum_{i=1}^N \theta_i(u_i) = 0 \ , \tag{4.2c}$$

where the index i can represent unit, time, uncertainties or a mix. Under mild technical conditions, this problem is equivalent to

$$\min_{\{u_i\}_{i=1}^N} \quad \max_{\lambda \in \mathbb{R}} \quad \sum_{i=1}^N L_i(u_i) + \lambda \Big(\sum_{i=1}^N \theta_i(u_i) \Big) \tag{4.3a}$$

$$s.t. \quad u_i \in U_i^{ad} \ , \qquad \forall i \in [\![1, N]\!] \ . \tag{4.3b}$$

Under a proper *constraint qualification condition*, we can exchange the min operator with the max operator and obtain

$$\max_{\lambda \in \mathbb{R}} \quad \min_{\{u_i\}_{i=1}^N} \quad \sum_{i=1}^N L_i(u_i) + \lambda \theta_i(u_i) \tag{4.4a}$$

$$s.t. \quad u_i \in U_i^{ad} \ , \qquad \forall i \in [\![1, N]\!] \ . \tag{4.4b}$$

Now, consider the inner minimization problem: the objective function is given as a sum of local costs, each of them determined by local independent controls. Thus, the minimum of

the sum is the sum of the minima, and Problem (4.4) can be written as

$$\max_{\lambda\in\mathbb{R}} \sum_{i=1}^{N} \min_{u_i} \quad L_i(u_i) + \lambda\theta_i(u_i) \tag{4.5a}$$

$$s.t. \quad u_i \in U_i^{ad} \ . \tag{4.5b}$$

For a given $\lambda = \lambda^{(k)}$, we now obtain N separate minimization problems, that are the subproblems of the decomposition method:

$$\min_{u_i} \quad L_i(u_i) + \lambda^{(k)}\theta_i(u_i) \tag{4.6a}$$

$$s.t. \quad u_i \in U_i^{ad} \ . \tag{4.6b}$$

These subproblems are updated as the multiplier $\lambda^{(k)}$ (or equivalently the price) is updated, like with

$$\lambda^{(k+1)} = \lambda^{(k)} + \rho \sum_{i=1}^{N} \theta_i(u_i^{(k)}) \ , \tag{4.7}$$

where $\rho > 0$ is a given parameter, and $u_i^{(k)}$ an optimal solution of Problem (4.6). This update formula for the multiplier is part of the equations of the dual gradient algorithm recalled next.

Remark 4.1 *The dual gradient algorithm (also known as the Uzawa Algorithm [25, 72]) can be presented as follows. We consider the following optimization problem*

$$\min_{u\in\mathcal{U}^{ad}} \quad J(u) \tag{4.8a}$$

$$s.t. \quad \theta(u) = 0 \ , \tag{4.8b}$$

where J is a a-strongly convex and Gâteau differentiable function, $\theta : \mathcal{U} \to \mathcal{V}$ is a linear, κ-Lipschitz operator between Hilbert spaces, $\mathcal{U}^{ad}$ is a closed convex subset of $\mathcal{U}$, and the Lagrangian $L(u, \lambda) = J(u) + \langle\lambda, \theta(u)\rangle$ admits a saddle-point (i.e. the constraints $\theta(u) = 0$ are qualified). Then the following algorithm converges toward the optimal solution of Problem (4.8a), when $0 < \rho < 2a/\kappa^2$.

Data: Initial multiplier λ_0, step ρ
Result: optimal control and multiplier;
repeat
  $u^{(k)} = \arg\min_{u\in\mathcal{U}^{ad}} L(u, \lambda^{(k)})$;
  $\lambda^{(k+1)} = \lambda^{(k)} + \rho\theta(u^{(k)})$;
until $\theta(u^{(k)}) = 0$;

Algorithm 4.1: Dual Gradient Algorithm

Remark 4.2 *This price decomposition scheme is the simplest and most well-known of decomposition schemes, but not the only one. In short, the* decomposition by quantity *approach consists in allocating to each subproblem a given quantity of the demand to satisfy, and then update the allocation; the* decomposition by prediction *approach consists in allocating to each subproblem a part of the constraint.*

Notice that, even if the property of having a sum of costs over units seems to be fundamental for decomposition, the Auxiliary Problem Principle (see [19]) allows to extend these decomposition schemes to general (non-additive) costs and constraint functions.

4.1.2 Spatial decomposition

The spatial decomposition (by prices) relies on the idea of dualizing the coupling constraint (4.1c).

We now apply to Problem (4.1) a price decomposition scheme, presented in §4.1.1, by dualizing the spatial constraint (4.1c). Since there are $T \times |\Omega|$ constraints of dimension n_c, the set of multipliers is of dimension $T \times |\Omega| \times n_c$. Problem (4.1), with constraint (4.1c) dualized, reads

$$\min_{\boldsymbol{X},\boldsymbol{U}} \quad \max_{\boldsymbol{\lambda}} \quad \sum_{\omega\in\Omega} \mathbb{P}\big(\{\omega\}\big) \sum_{t=0}^{T-1} \Bigg(\sum_{i=1}^{N} L_t^i\Big(\boldsymbol{X}_t^i(\omega), \boldsymbol{U}_t^i(\omega), \boldsymbol{W}_t(\omega)\Big) \tag{4.9a}$$

$$+ \boldsymbol{\lambda}_t(\omega) \cdot \sum_{i=1}^{N} \theta_t^i\big(\boldsymbol{X}_t^i(\omega), \boldsymbol{U}_t^i(\omega)\big)\Bigg) \tag{4.9b}$$

$$s.t. \quad \boldsymbol{X}_{t+1}^i(\omega) = F_t^i\big(\boldsymbol{X}_t^i(\omega), \boldsymbol{U}_t^i(\omega), \boldsymbol{W}_t(\omega)\big) \qquad \forall t\ ,\ \forall i\ ,\ \forall \omega\ , \tag{4.9c}$$

$$\boldsymbol{U}_t^i \preceq \mathcal{F}_t\ , \qquad \forall t\ ,\ \forall i\ . \tag{4.9d}$$

Assuming constraint qualification, this problem is equivalent to

$$\max_{\boldsymbol{\lambda}\in\Lambda} \quad \sum_{i=1}^{N} \quad \min_{\boldsymbol{X}^i,\boldsymbol{U}^i} \quad \sum_{\omega\in\Omega} \mathbb{P}\big(\{\omega\}\big) \sum_{t=0}^{T-1} \quad L_t^i\Big(\boldsymbol{X}_t^i(\omega), \boldsymbol{U}_t^i(\omega), \boldsymbol{W}_t(\omega)\Big) \tag{4.10a}$$

$$+ \boldsymbol{\lambda}_t(\omega) \cdot \theta_t^i\big(\boldsymbol{X}_t^i(\omega), \boldsymbol{U}_t^i(\omega)\big) \tag{4.10b}$$

$$s.t. \quad \boldsymbol{X}_{t+1}^i(\omega) = F_t^i\big(\boldsymbol{X}_t^i(\omega), \boldsymbol{U}_t^i(\omega), \boldsymbol{W}_t(\omega)\big) \quad \forall t\ ,\ \forall i\ ,\ \forall \omega\ , \tag{4.10c}$$

$$\boldsymbol{U}_t^i \preceq \mathcal{F}_t\ , \qquad \forall t\ ,\ \forall i\ . \tag{4.10d}$$

For a given multiplier $\boldsymbol{\lambda}^{(k)}$, we obtain N parallel inner minimization problems

$$\min_{\boldsymbol{X}^i,\boldsymbol{U}^i} \quad \sum_{\omega\in\Omega} \mathbb{P}\big(\{\omega\}\big) \sum_{t=0}^{T-1} \quad L_t^i\Big(\boldsymbol{X}_t^i(\omega), \boldsymbol{U}_t^i(\omega), \boldsymbol{W}_t(\omega)\Big) \tag{4.11a}$$

$$+ \boldsymbol{\lambda}_t^{(k)}(\omega) \cdot \theta_t^i\big(\boldsymbol{X}_t^i(\omega), \boldsymbol{U}_t^i(\omega)\big) \tag{4.11b}$$

$$s.t. \quad \boldsymbol{X}_{t+1}^i(\omega) = F_t^i\big(\boldsymbol{X}_t^i(\omega), \boldsymbol{U}_t^i(\omega), \boldsymbol{W}_t(\omega)\big) \qquad \forall t\ ,\ \forall \omega \tag{4.11c}$$

$$\boldsymbol{U}_t^i \preceq \mathcal{F}_t\ , \qquad \forall t\ . \tag{4.11d}$$

We denote $\boldsymbol{U}_t^{i,(k)}$ and $\boldsymbol{X}_t^{i,(k)}$ an optimal solution. We update the multipliers with

$$\boldsymbol{\lambda}_t^{(k+1)}(\omega) = \boldsymbol{\lambda}_t^{(k)}(\omega) + \rho\Big(\sum_{i=1}^{N} \theta_t^i\Big(\boldsymbol{X}_t^{(k)}(\omega), \boldsymbol{U}_t^{(k)}(\omega)\Big)\Big) \ , \quad \forall t \in [\![0, T-1]\!] \ , \quad \forall \omega \in \Omega \ , \tag{4.12}$$

where $\rho > 0$ is a given parameter.

Remark 4.3 *As discussed in §4.1.1, this price decomposition has an insightful interpretation. The multiplier $\boldsymbol{\lambda}_t(\omega)$ can be interpreted as the marginal cost of the output at time t along scenario ω. It is worth noting that the prices form a stochastic process $\{\boldsymbol{\lambda}_t\}_{t=0}^{T-1}$, that can be represented as an element of the huge space $\mathbb{R}^{(Tn_c)^{|\Omega|}}$. We can show that we can only consider non-anticipative processes. The DADP method presented in §5.8 consists precisely in restricting the space of multipliers $\boldsymbol{\lambda}$ over which the maximization is done.*

4.1.3 Scenario decomposition

The decomposition scenario by scenario consists in dualizing the non-anticipativity constraint, and then solving subproblems for each scenario (using any of the tools available for deterministic problems). The Progressive Hedging (PH) Algorithm stands as the state of the art in this domain, but we also present the Stochastic Pontryaguin approach.

Principle of scenario decomposition

We consider Problem (4.1) written on a tree $\mathcal{T}$. We then have

$$\min_{\left\{\{u_{n_t}\}_{n_t \in \mathcal{N}_t}\right\}_{t=0}^{T-1}} \quad \sum_{n \in \mathcal{T}} \sum_{m \in c(n)} \sum_{i=1}^{N} \mathbb{P}\big(\{m\}\big)\ L_t^i\Big(x_n^i, u_m^i, w_m\Big) \tag{4.13a}$$

$$s.t. \quad x_m^i = F_t^i\big(x_n^i, u_m^i, w_m\big) \ , \qquad \forall i \ , \ \forall m \in c(n) \ , \ \forall n \in \mathcal{N}_t \ , \ \forall t, \tag{4.13b}$$

$$\sum_{i=1}^{N} \theta_t^i\big(x_n^i, u_m^i\big) = 0 \ , \qquad \forall i \ , \ \forall m \in c(n) \ , \ \forall n \in \mathcal{N}_t \ , \ \forall t \ . \tag{4.13c}$$

Note that we have one decision u_n per node on the tree; this materializes the information constraint (4.1d), the one that is dualized in scenario decomposition algorithms (like Progressive Hedging). For this purpose, we re-introduce control variables (see Figure 3.2), that is, a sequence $\{\boldsymbol{U}_t(\omega)\}_{t=0}^{T-1}$ of controls for each scenario ω (associated to a leaf of the tree), as in Problem (4.1). It means that, with a given node $n \in \mathcal{T}$, are associated $|n|$ control variables, that is, one per scenario going through this node. The non-anticipativity constraint (4.1d) is represented by

$$\boldsymbol{U}_t(\omega) = \boldsymbol{U}_t(\omega') \ , \quad \forall t \in [\![0, T-1]\!] \ , \quad \forall (\omega, \omega') \in n^2, \ . \tag{4.14}$$

We introduce $\overline{U}_n$, the mean control on node $n \in \mathcal{N}_t$, defined by

$$\overline{\boldsymbol{U}}_n = \frac{\sum_{\omega \in n} \boldsymbol{U}_t(\omega)}{|n|} \ . \tag{4.15}$$

We denote by $n_t(\omega)$ the node of depth t to which the scenario ω belongs. Hence, Equation (4.14) can be rewritten as

$$\boldsymbol{U}_t(\omega) = \overline{\boldsymbol{U}}_{n_t(\omega)} \ , \ \ \forall t \in [\![0, T-1]\!] \ , \ \ \forall \omega \in \Omega \ , \tag{4.16}$$

and Problem (4.1) now reads

$$\min_{\boldsymbol{X}, \boldsymbol{U}} \quad \sum_{\omega \in \Omega} \mathbb{P}\big(\{\omega\}\big) \sum_{i=1}^{N} \sum_{t=0}^{T-1} L_t^i\Big(\boldsymbol{X}_t^i(\omega), \boldsymbol{U}_t^i(\omega), \boldsymbol{W}_t(\omega)\Big) \tag{4.17a}$$

$$s.t. \quad \boldsymbol{X}_{t+1}^i(\omega) = F_t^i\big(\boldsymbol{X}_t^i(\omega), \boldsymbol{U}_t^i(\omega), \boldsymbol{W}_t(\omega)\big) \ , \qquad \forall t \ , \ \ \forall i \ , \ \ \forall \omega \tag{4.17b}$$

$$\sum_{i=1}^{N} \theta_t^i\big(\boldsymbol{X}_t^i(\omega), \boldsymbol{U}_t^i(\omega)\big) = 0 \ , \qquad \forall t, \quad \forall \omega \tag{4.17c}$$

$$\boldsymbol{U}_t(\omega) = \overline{\boldsymbol{U}}_{n_t(\omega)} \qquad \forall t \ , \ \ \forall \omega \ . \tag{4.17d}$$

Now, when we dualize Constraint (4.17d) under proper constraint qualification conditions, we obtain

$$\max_{\boldsymbol{\lambda} \in \Lambda} \ \min_{\boldsymbol{X}, \boldsymbol{U}} \quad \sum_{\omega \in \Omega} \mathbb{P}\big(\{\omega\}\big) \sum_{i=1}^{N} \sum_{t=0}^{T-1} L_t^i\Big(\boldsymbol{X}_t^i(\omega), \boldsymbol{U}_t^i(\omega), \boldsymbol{W}_t(\omega)\Big) \tag{4.18a}$$

$$+ \boldsymbol{\lambda}_t(\omega)\big(\boldsymbol{U}_t(\omega) - \overline{\boldsymbol{U}}_{n_t(\omega)}\big) \tag{4.18b}$$

$$s.t. \quad \boldsymbol{X}_{t+1}^i(\omega) = F_t^i\big(\boldsymbol{X}_t^i(\omega), \boldsymbol{U}_t^i(\omega), \boldsymbol{W}_t(\omega)\big) \ , \qquad \forall t \ , \ \ \forall i \ , \ \ \forall \omega \tag{4.18c}$$

$$\sum_{i=1}^{N} \theta_t^i\big(\boldsymbol{X}_t^i(\omega), \boldsymbol{U}_t^i(\omega)\big) = 0 \ , \qquad \forall t, \quad \forall \omega \ , \tag{4.18d}$$

where Λ is of dimension $|\Omega| \times N \times n_u \times T$.

In this last formulation, the term $\overline{\boldsymbol{U}}_{n_t(\omega)}$ is not a new variable but a linear operator, as expressed by Equation (4.20). As a consequence, the term

$$\sum_{\omega \in \Omega} \mathbb{P}\big(\{\omega\}\big) \boldsymbol{\lambda}_t(\omega) \overline{\boldsymbol{U}}_{n_t(\omega)} = \sum_{\omega \in \Omega} \mathbb{P}\big(\{\omega\}\big) \boldsymbol{\lambda}_t(\omega) \frac{1}{|n_t(\omega)|} \sum_{\omega' \in n_t(\omega)} \boldsymbol{U}_t(\omega') \tag{4.19}$$

induces a coupling $\omega' \in n_t(\omega)$ between scenarios.

First approach to decoupling. As the non-anticipativity constraint (4.16) is equivalent to

$$\boldsymbol{U}_t^i - \mathbb{E}\big[\boldsymbol{U}_t^i \;\big|\; \mathcal{F}_t\big] = 0\ ,\ \ \forall i \in [\![1, N]\!]\ ,\ \ \forall t \in [\![0, T-1]\!]\ , \tag{4.20}$$

we can remove this coupling by using the adjoint operator of the operator $I_d - \mathbb{E}\big[\,\cdot\;\big|\;\mathcal{F}_t\big]$ as done in [56]. This leads to the decoupled formulation

$$\max_{\boldsymbol{\lambda}\in\Lambda'}\ \min_{\boldsymbol{X},\boldsymbol{U}}\quad \sum_{\omega\in\Omega}\mathbb{P}\big(\{\omega\}\big)\sum_{i=1}^{N}\sum_{t=0}^{T-1} L_t^i\Big(\boldsymbol{X}_t^i(\omega), \boldsymbol{U}_t^i(\omega), \boldsymbol{W}_t(\omega)\Big) \tag{4.21a}$$

$$+\,\boldsymbol{\lambda}_t(\omega)\boldsymbol{U}_t(\omega) \tag{4.21b}$$

$$s.t.\quad \boldsymbol{X}_{t+1}^i(\omega) = F_t^i\big(\boldsymbol{X}_t^i(\omega), \boldsymbol{U}_t^i(\omega), \boldsymbol{W}_t(\omega)\big)\ ,\qquad \forall t\ ,\ \ \forall i\ ,\ \ \forall\omega \tag{4.21c}$$

$$\sum_{i=1}^{N}\theta_t^i\big(\boldsymbol{X}_t^i(\omega), \boldsymbol{U}_t^i(\omega)\big) = 0\ , \qquad \forall t,\quad \forall\omega\ , \tag{4.21d}$$

but where the multipliers are restricted to the subset $\Lambda' \subset \Lambda$, containing multipliers $\boldsymbol{\lambda}$ which satisfy $\sum_{\omega\in n}\boldsymbol{\lambda}_t(\omega) = 0$, for all node $n \in \mathcal{N}_t$.

Second approach to decoupling. To achieve scenarios decoupling, we can also use variable duplication. Here, this consists in considering that the variables $\overline{U}_{n_t(\omega)}$ are new minimizing variables. We now take this road to describe the Progressive Hedging algorithm. For this purpose, we reformulate Problem (4.18) as a special case, when $r = 0$, of

$$\max_{\boldsymbol{\lambda}\in\Lambda}\ \min_{\boldsymbol{X},\boldsymbol{U},\overline{U}}\quad \sum_{\omega\in\Omega}\mathbb{P}\big(\{\omega\}\big)\sum_{i=1}^{N}\sum_{t=0}^{T-1} L_t^i\Big(\boldsymbol{X}_t^i(\omega), \boldsymbol{U}_t^i(\omega), \boldsymbol{W}_t(\omega)\Big) \tag{4.22a}$$

$$+\,\boldsymbol{\lambda}_t^{(k)}(\omega)\big(\boldsymbol{U}_t(\omega) - \overline{\boldsymbol{U}}_{n_t(\omega)}\big) + \frac{r}{2}\Big\|\boldsymbol{U}_t(\omega) - \overline{\boldsymbol{U}}^{(k)}_{n_t(\omega)}\Big\|^2 \tag{4.22b}$$

$$s.t.\quad \boldsymbol{X}_{t+1}^i(\omega) = F_t^i\big(\boldsymbol{X}_t^i(\omega), \boldsymbol{U}_t^i(\omega), \boldsymbol{W}_t(\omega)\big)\ , \qquad \forall t\ ,\ \ \forall i\ ,\ \ \forall\omega \tag{4.22c}$$

$$\sum_{i=1}^{N}\theta_t^i\big(\boldsymbol{X}_t^i(\omega), \boldsymbol{U}_t^i(\omega)\big) = 0\ , \qquad \forall t,\quad \forall\omega\ , \tag{4.22d}$$

Hence, we have now inserted an augmented Lagrangian term, which is active when $r > 0$. The Progressive Hedging algorithm proposes a decomposition algorithm as follows.

Progressive Hedging (PH)

The inner step at iteration k consists in the following problem. For each scenario $\omega \in \Omega$, and each stage $t \in [\![0, T-1]\!]$, a multiplier $\boldsymbol{\lambda}_t^{(k)}(\omega)$ is fixed, and the inner minimization problem to solve is:

$$\min_{\boldsymbol{X}(\omega),\boldsymbol{U}(\omega),\overline{\boldsymbol{U}}_{n_t(\omega)}} \quad \sum_{i=1}^{N}\sum_{t=0}^{T-1} L_t^i\Big(\boldsymbol{X}_t^i(\omega),\boldsymbol{U}_t^i(\omega),\boldsymbol{W}_t(\omega)\Big) \tag{4.23a}$$

$$+\boldsymbol{\lambda}_t^{(k)}(\omega)\big(\boldsymbol{U}_t(\omega)-\overline{\boldsymbol{U}}_{n_t(\omega)}\big)+\frac{r}{2}\Big\|\boldsymbol{U}_t(\omega)-\overline{\boldsymbol{U}}_{n_t(\omega)}^{(k)}\Big\|^2 \tag{4.23b}$$

$$s.t. \quad \boldsymbol{X}_{t+1}^i(\omega)=F_t^i\big(\boldsymbol{X}_t^i(\omega),\boldsymbol{U}_t^i(\omega),\boldsymbol{W}_t(\omega)\big)\ , \qquad \forall t\ ,\ \forall i\ ,\ \forall\omega \tag{4.23c}$$

$$\sum_{i=1}^{N}\theta_t^i\big(\boldsymbol{X}_t^i(\omega),\boldsymbol{U}_t^i(\omega)\big)=0\ , \qquad \forall t,\quad \forall\omega\ . \tag{4.23d}$$

We obtain scenario decoupling by using an alternate minimization step, as described in Algorithm 4.2. When $\overline{\boldsymbol{U}}$ is fixed, the minimization with respect to $\boldsymbol{U}$ achieves scenario decomposition. When $\boldsymbol{U}$ is fixed, the minimization with respect to $\overline{\boldsymbol{U}}$ is done explicitly, leading to

$$\overline{\boldsymbol{U}}_n^{(k+1)}=\frac{\sum_{\omega\in n}\boldsymbol{U}_t^{(k+1)}(\omega)}{|n|}\ ,\ \ \forall t\in[\![0,T-1]\!]\ ,\ \ \forall n\in\mathcal{N}_t\ . \tag{4.24}$$

The general scheme of the Progressive Hedging algorithm is described in Algorithm 4.2.

Data: Initial multipliers $\big\{\{\boldsymbol{\lambda}_t^{(0)}(\omega)\}_{t=0}^{T-1}\big\}_{\omega\in\Omega}$ and mean control $\big\{\overline{\boldsymbol{U}}_n^{(0)}\big\}_{n\in\mathcal{T}}$;
Result: optimal feedback;
repeat
  forall the *scenario* $\omega\in\Omega$ **do**
    Solves the deterministic minimization problem (4.23) for scenario ω with a measurability penalization, and obtain optimal control $\boldsymbol{U}^{(k+1)}$;
  Update the mean controls

$$\overline{\boldsymbol{U}}_n^{(k+1)}=\frac{\sum_{\omega\in n}\boldsymbol{U}_t^{(k+1)}(\omega)}{|n|}\ ,\ \ \forall t\in[\![0,T-1]\!]\ ,\ \ \forall n\in\mathcal{N}_t\ ; \tag{4.25}$$

  Update the measurability penalization with

$$\boldsymbol{\lambda}_t^{(k+1)}(\omega)=\boldsymbol{\lambda}_t^{(k)}(\omega)+\rho\big(\boldsymbol{U}_t(\omega)^{(k+1)}-\overline{\boldsymbol{U}}_{n_t(\omega)}^{(k+1)}\big)\ ,\ \ \forall\omega\in\Omega\ ,\ \ \forall t\in[\![0,T-1]\!]\ . \tag{4.26}$$

until $\boldsymbol{U}_t^i-\mathbb{E}\big[\boldsymbol{U}_t^i\ \big|\ \mathcal{F}_t\big]=0$;

Algorithm 4.2: General Scheme of Scenario Decomposition

The scenario decomposition principle that we just exposed is an illustration of the price decomposition scheme outlined in §4.1.1. Gradient steps for the multipliers $\boldsymbol{\lambda}_t^{(k)}(\omega)$ in Algorithm 4.2 mimick those in equation (4.7), in order to maximize the dual function with respect to the multiplier λ in (4.22).

However, in Algorithm 4.2 the dual function is regularized using augmented Lagrangian ($r > 0$). It is well known that using an augmented Lagrangian, instead of a simple Lagrangian, can enlarge the set of problems which can be solved by maximizing the dual function. For example, strong convexity assumption for the criterion can be dropped.

As we have seen, using an augmented Lagrangian introduces a quadratic term which breaks the scenario by scenario decoupling. Then, PH uses the special form of the dualized constraint in (4.23b) to recover separability scenario by scenario. A two step minimization is used, one with respect to the variables $\boldsymbol{U}_t^i$ for fixed values of the mean controls $\overline{\boldsymbol{U}}_n^{(k)}$ which can be done scenario by scenario; and one step were the minimization with respect to the mean control is performed analytically. We refer the reader to [56, 16] for more details.

Note also that maximizing the dual function in the augmented Lagrangian setting is related to the proximal Algorithm in the convex case and equality constraints.

Stochastic Pontryaguin

We present an extension to the stochastic framework of Pontryaguin method. More details and numerical experiments can be found in [20].

Ignoring the "spatial" coupling constraint (4.1c), and dualizing the dynamics constraints (4.1b), Problem (4.1) reads

$$\min_{\{\boldsymbol{U}_t \preceq \mathcal{F}_t\}_{t=0}^{T-1}} \left\{ \min_{\boldsymbol{X}} \max_{\boldsymbol{\lambda}\in\Lambda} \sum_{\omega\in\Omega} \sum_{i=1}^{N} \sum_{t=0}^{T-1} \mathbb{P}\big(\{\omega\}\big)\, L_t^i\Big(\boldsymbol{X}_t^i(\omega), \boldsymbol{U}_t^i(\omega), \boldsymbol{W}_t(\omega)\Big) + \boldsymbol{\lambda}_{t+1}^i(\omega)\Big(F_t^i\big(\boldsymbol{X}_t^i(\omega), \boldsymbol{U}_t^i(\omega), \boldsymbol{W}_t(\omega)\big) - \boldsymbol{X}_{t+1}^i(\omega)\Big) \right\}. \tag{4.27}$$

For a given control process $\boldsymbol{U}^{(k)}$, we consider the inner min-max problem,

$$\min_{\boldsymbol{X}} \max_{\boldsymbol{\lambda}} \sum_{\omega\in\Omega} \mathbb{P}\big(\{\omega\}\big) \sum_{i=1}^{N} \sum_{t=0}^{T-1} L_t^i\Big(\boldsymbol{X}_t^i(\omega), \boldsymbol{U}_t^{i,(k)}(\omega), \boldsymbol{W}_t(\omega)\Big) + \boldsymbol{\lambda}_{t+1}^i(\omega)\Big(F_t^i\big(\boldsymbol{X}_t^i(\omega), \boldsymbol{U}_t^{i,(k)}(\omega), \boldsymbol{W}_t(\omega)\big) - \boldsymbol{X}_{t+1}^i(\omega)\Big). \tag{4.28}$$

This problem can be solved ω per ω

$$\min_{\boldsymbol{X}(\omega)} \max_{\boldsymbol{\lambda}(\omega)} \sum_{i=1}^{N} \sum_{t=0}^{T-1} L_t^i\Big(\boldsymbol{X}_t^i(\omega), \boldsymbol{U}_t^{i,(k)}(\omega), \boldsymbol{W}_t(\omega)\Big) + \boldsymbol{\lambda}_{t+1}^i(\omega)\Big(F_t^i\big(\boldsymbol{X}_t^i(\omega), \boldsymbol{U}_t^{i,(k)}(\omega), \boldsymbol{W}_t(\omega)\big) - \boldsymbol{X}_{t+1}^i(\omega)\Big). \tag{4.29}$$

Assuming that we have the necessary regularity conditions (and since we assumed no bound constraints on $\boldsymbol{U}$ and $\boldsymbol{X}$), we write the first order optimality conditions of this inner

min-max problem and deduce the optimal solutions $\boldsymbol{X}^{(k)}$ and $\boldsymbol{\lambda}^{(k)}$ by

$$\boldsymbol{X}_0^{(k)} = x_0 \ , \tag{4.30a}$$
$$\boldsymbol{X}_{t+1}^{(k)} = F_t\big(\boldsymbol{X}_t^{(k)}, \boldsymbol{U}_t^{(k)}, \boldsymbol{W}_t\big) \qquad t \in [\![0, T-1]\!] \ , \tag{4.30b}$$
$$\boldsymbol{\lambda}_T = 0 \ , \tag{4.30c}$$
$$\boldsymbol{\lambda}_t = \nabla_x F_t\big(\boldsymbol{X}_t^{(k)}, \boldsymbol{U}_t^{(k)}, \boldsymbol{W}_t\big)\boldsymbol{\lambda}_{t+1}^{(k)} + \nabla_x L_t\big(\boldsymbol{X}_t^{(k)}, \boldsymbol{U}_t^{(k)}, \boldsymbol{W}_t\big) \qquad t \in [\![1, T-1]\!] \ . \tag{4.30d}$$

These conditions involve a co-state stochastic process $\boldsymbol{\lambda}$ which is *not* $\mathcal{F}$-adapted since the dynamics (4.30c)–(4.30d) propagate backwards and therefore $\boldsymbol{\lambda}_t$ is not $\mathcal{F}_t$-measurable in general.

Given a control trajectory $(\boldsymbol{U}_0^{(k)}, \dots, \boldsymbol{U}_{T-1}^{(k)})$, we can solve these equations by, first integrating Equations (4.30a)-(4.30b) forward to obtain $\{\boldsymbol{X}_t^{(k)}\}_{t=0}^T$, and then integrating Equations (4.30c)-(4.30d) backward to obtain the multiplier process $\{\boldsymbol{\lambda}_t^{(k)}\}_{t=1}^T$. Note that these integrations are performed scenario per scenario, hence in parallel.

Denote by H the function mini-maximized in Problem (4.27), i.e.

$$H\big(\boldsymbol{X}, \boldsymbol{U}, \boldsymbol{\lambda}\big) = \sum_{\omega\in\Omega} \mathbb{P}\big(\{\omega\}\big) \sum_{i=1}^{N} \sum_{t=0}^{T-1} \quad L_t^i\Big(\boldsymbol{X}_t^i(\omega), \boldsymbol{U}_t^i(\omega), \boldsymbol{W}_t(\omega)\Big) \tag{4.31}$$
$$+ \boldsymbol{\lambda}_{t+1}^i(\omega)\Big(F_t^i\big(\boldsymbol{X}_t^i(\omega), \boldsymbol{U}_t^i(\omega), \boldsymbol{W}_t(\omega)\big) - \boldsymbol{X}_{t+1}^i(\omega)\Big) \ .$$

Define by J the function minimized in Problem (4.27), that is,

$$J(\boldsymbol{U}) = \min_{\boldsymbol{X}} \quad \max_{\boldsymbol{\lambda}} \quad H\big(\boldsymbol{X}, \boldsymbol{U}, \boldsymbol{\lambda}\big) \ . \tag{4.32}$$

The Danskin theorem (also known as the *envelop theorem* in Economics), states that, under proper assumptions, the gradient of the function J at point $\boldsymbol{U}^{(k)}$ is given by

$$\nabla J(\boldsymbol{U}^{(k)}) = \nabla_{\boldsymbol{U}} H(\boldsymbol{X}^{(k)}, \boldsymbol{U}^{(k)}, \boldsymbol{\lambda}^{(k)}) \ . \tag{4.33}$$

Hence, the gradient of J at $\boldsymbol{U}^{(k)}$ is

$$\nabla J(\boldsymbol{U}^{(k)}) = \nabla_u L_t\big(\boldsymbol{X}_t^{(k)}, \boldsymbol{U}_t^{(k)}, \boldsymbol{W}_t\big) + \nabla_u F_t\big(\boldsymbol{X}_t^{(k)}, \boldsymbol{U}_t^{(k)}, \boldsymbol{W}_t\big)\boldsymbol{\lambda}_{t+1}^{(k)} \ . \tag{4.34}$$

As the minimization is done over the adapted controls, a projected gradient step for the minimization of J would be

$$\boldsymbol{U}_t^{(k+1)} = \boldsymbol{U}_t^{(k)} + \rho\mathbb{E}\Big[\nabla_u L_t\big(\boldsymbol{X}_t^{(k)}, \boldsymbol{U}_t^{(k)}, \boldsymbol{W}_t\big) + \nabla_u F_t\big(\boldsymbol{X}_t^{(k)}, \boldsymbol{U}_t^{(k)}, \boldsymbol{W}_t\big)\boldsymbol{\lambda}_{t+1}^{(k)} \ \Big| \ \mathcal{F}_t\Big] \ . \tag{4.35}$$

Equation (4.35) can be used as an update step of the control $\boldsymbol{U}_t^{(k)}$ for this decomposition method.

4.1.4 Time decomposition

Not all decompositions by duality lead to powerful formulations. For instance, we present a (little used) parallel decomposition approach of time decomposition obtained by dualization of the dynamic constraint.

On the other hand, as there is a natural flow in time, we can write a chained decomposition method, the well-known *Dynamic Programming* approach.

Dualizing the dynamics constraints

We apply to Problem (4.1) a price decomposition scheme, presented in §4.1.1, by dualizing the dynamic constraint (4.1b).

Since there are $N \times T \times |\Omega|$ dynamics constraints, the set of multiplier is of dimension $T \times |\Omega| \times N \times n_X$. Dualizing the dynamics constraints (4.1b), Problem (4.1) reads

$$\begin{aligned}
\min_{\boldsymbol{X},\boldsymbol{U}} \quad \max_{\boldsymbol{\lambda}} \quad & \sum_{\omega\in\Omega}\sum_{i=1}^{N}\sum_{t=0}^{T-1} \mathbb{P}\big(\{\omega\}\big)\, L_t^i\Big(\boldsymbol{X}_t^i(\omega), \boldsymbol{U}_t^i(\omega), \boldsymbol{W}_t(\omega)\Big) \\
& \qquad + \boldsymbol{\lambda}_t^i(\omega)\boldsymbol{X}_t^i(\omega) - \boldsymbol{\lambda}_{t+1}^i(\omega) F_t^i\big(\boldsymbol{X}_t^i(\omega), \boldsymbol{U}_t^i(\omega), \boldsymbol{W}_t(\omega)\big)\ , \\
s.t. \quad & \sum_{i=1}^{N} \theta_t^i\big(\boldsymbol{X}_t^i(\omega), \boldsymbol{U}_t^i(\omega)\big) = 0\ , && \forall t, \quad \forall \omega \\
& \boldsymbol{U}_t^i \preceq \mathcal{F}_t && \forall t\ , \ \forall i\ .
\end{aligned}$$

Assuming constraint qualification, and fixing a multiplier $\boldsymbol{\lambda}^{(k)}$, we obtain T separate inner minimization problems

$$\begin{aligned}
\min_{\boldsymbol{X}_t,\boldsymbol{U}_t} \quad & \sum_{\omega\in\Omega}\sum_{i=1}^{N} \mathbb{P}\big(\{\omega\}\big)\, L_t^i\Big(\boldsymbol{X}_t^i(\omega), \boldsymbol{U}_t^i(\omega), \boldsymbol{W}_t(\omega)\Big) \\
& \qquad + \boldsymbol{\lambda}_t^{i,(k)}(\omega)\boldsymbol{X}_t^i(\omega) - \boldsymbol{\lambda}_{t+1}^{i,(k)}(\omega) F_t^i\big(\boldsymbol{X}_t^i(\omega), \boldsymbol{U}_t^i(\omega), \boldsymbol{W}_t(\omega)\big)\ , \\
s.t. \quad & \sum_{i=1}^{N} \theta_t^i\big(\boldsymbol{X}_t^i(\omega), \boldsymbol{U}_t^i(\omega)\big) = 0\ , && \forall \omega \\
& \boldsymbol{U}_t^i \preceq \mathcal{F}_t && \forall i\ .
\end{aligned}$$

We denote $\boldsymbol{U}_t^{i,(k)}$ and $\boldsymbol{X}_t^{i,(k)}$ an optimal solution. We update the multipliers with

$$\boldsymbol{\lambda}_{t+1}^{i,(k+1)}(\omega) = \boldsymbol{\lambda}_{t+1}^{i,(k)}(\omega) + \rho\Big(\boldsymbol{X}_{t+1}^{i,(k)}(\omega) - F_t^i\big(\boldsymbol{X}_t^{i,(k)}(\omega), \boldsymbol{U}_t^{i,(k)}(\omega), \boldsymbol{W}_t(\omega)\big)\Big)\ . \tag{4.36}$$

This decomposition approach is probably one of the less used decomposition approaches.

Dynamic Programming (DP)

The Dynamic Programming method is a well-known decomposition in time (see [5]). As it is usual, we present the Dynamic Programming in a Decision-Hazard setting. It relies on

the assumption that the exogenous noises $\{\boldsymbol{W}_t\}_0^{T-1}$ form a sequence of independent random variables. With this assumption, the original state $\boldsymbol{X} = (\boldsymbol{X}_1, \dots, \boldsymbol{X}_N)$, that follows (4.1b), is a so-called *information state* (see Remark 3.2). This state is the argument of the *value function* V_t: $V_t(x)$ is the best possible future cost starting from time t in state x. The value functions satisfy the *Dynamic Programming Equations*: the V_t are computed backwards, starting from V_T and solving static optimization problems (see Algorithm 4.3). The solutions of these static optimization problems provide an optimal solution as a deterministic function of the current state (state feedback)

$$\boldsymbol{U}_t^\star = \pi_t^\star\big(\boldsymbol{X}_t^\star\big) \ , \tag{4.37}$$

where $\pi_t^\star : \mathbb{X}_t \to \mathbb{U}_t$. Observe that the solution, supposed to satisfy the non-anticipativity constraint (4.1d), satisfies what is a stronger constraint, namely $\boldsymbol{U}_t \preceq \boldsymbol{X}_t$. This is an important property of DP: when the exogenous noises $\{\boldsymbol{W}_t\}_{t=0}^{T-1}$ form a sequence of independent random variables, there is no loss of optimality in reducing the search to the class of state feedback solutions, namely $\mathcal{B}_t = \sigma(\boldsymbol{X}_t)$ instead of $\mathcal{B}_t = \mathcal{F}_t$. When the size of the state space $\mathbb{X}_t$ does not increase with t, this property has major consequences for numerical applications: whereas the space of $\mathcal{F}_t$-measurable solutions increases (exponentially) with t, the space of policies $\pi_t : \mathbb{X}_t \to \mathbb{U}_t$ does not.

Data: Problem data (especially initial point x_0 and final cost functions K^i);
Result: Bellman function V_t, Optimal feedback π_t ;
$V_T(x) = 0$;
for $t = T-1$ **to** 0 **do**
  foreach $x_t \in \mathbb{X}_t$ **do**

$$\begin{aligned} V_t(x_t) \quad = \quad & \min_{u=\{u^i\}_1^N} \quad \mathbb{E}\Big[\sum_{i=1}^N \quad L_t^i\big(x_t^i, u^i, \boldsymbol{W}_t\big) + V_{t+1}\big(\boldsymbol{X}_{t+1}\big)\Big] \\ & s.t. \quad \boldsymbol{X}_{t+1}^i = F_t^i\big(x_t^i, u^i, \boldsymbol{W}_t\big) \ , \ \ \forall i \\ & \qquad \sum_{i=1}^N \theta_t^i\big(x_t^i, u_t^i\big) = 0 \end{aligned} \tag{4.38}$$

  $\pi_t(x_t)$ is a control u minimizing the above problem ;

Algorithm 4.3: Dynamic Programming Algorithm

The DP chained decomposition is possible because of a causality principle along the time axis (this would not be possible for the uncertainty or for space, except under very specific conditions).

Remark 4.4 *Here, we make the major assumption that the size of the state space $\mathbb{X}_t$ does not increase with t. We suppose that each component of the state takes a finite number n_x of*

values (hence the state takes at most $\big(n_x\big)^N$ *values). Solving (4.1) by DP requires to explore*

$$T\big(n_x\big)^N n_u \tag{4.39}$$

possible solutions. Comparing with (3.38), we see that DP makes better than brute force whenever

$$\log T + N \log n_x + \log n_u \leq \frac{(n_w)^T - 1}{n_w - 1} \log n_u \ . \tag{4.40}$$

Therefore, the DP algorithm outbeats brute force for a large enough number T *of time steps. Indeed, it is linear in time, whereas brute force is exponential in time. However, the complexity of DP is exponential in the number* N *of subproblems or, in other words, in the dimension of the state: this stands as the* curse of dimensionality *(see [6]).*

Discussing DP and the notion of state

When the exogenous noises $\{\boldsymbol{W}_t\}_{t=0}^{T-1}$ form a sequence of independent random variables, we can write a *Dynamic Programming Equation* (DPE) like (4.38) with state $\boldsymbol{X}$. Now, what happens if this assumption fails? We lay out a theoretical and a practical answer.

The theoretical answer follows [79]. We introduce the "maximal" state (already mentioned in Table 3.3)

$$\widehat{\boldsymbol{X}}_t = \big(x_0, \boldsymbol{W}, \boldsymbol{U}_0, \dots, \boldsymbol{U}_{t-1}\big) \ , \tag{4.41}$$

which satisfies the trivial dynamic equation

$$\widehat{\boldsymbol{X}}_{t+1} = \big(\widehat{\boldsymbol{X}}_t, \boldsymbol{U}_t\big) \ . \tag{4.42}$$

Then, there exists a DPE, but with an even larger information state consisting of the conditional distribution of $\widehat{\boldsymbol{X}}_t$ knowing $\mathcal{F}_t$ [79]. Of course, this state is only of theoretical interest.

The practical answer has much to do with the "art of modelling", a compromise between, on the one hand, realism and complexity, and, on the other hand, mathematical tractability. Suppose you manage a dam. The natural physical state is the level of water in the dam, whereas the information state depends on the water inflows (rain, snow melting, etc.). To account for (a weak form of) dependency, we can make the assumption that the inflows are independent random variables, but that their distributions are not stationary, and depend upon time t to reflect seasonal effects. In that case, the physical state is an information state. To account for (a stronger form of) dependency, we can make the assumption that the inflows follow a so-called "order 1 model" (e.g. an AR-1 model)

$$\boldsymbol{W}_{t+1} = \widetilde{f}_t(\boldsymbol{W}_t, \widehat{\boldsymbol{W}}_t) \ , \tag{4.43}$$

where $\{\widehat{\boldsymbol{W}}_t\}_{t=0}^{T-1}$ is a sequence of independent random variables. Here, an information state is given by

$$\widehat{\boldsymbol{X}}_t = \big(\boldsymbol{X}_t, \boldsymbol{W}_t\big) \ , \tag{4.44}$$

with the dynamic

$$\widehat{\boldsymbol{X}}_{t+1} = \Big(F_t(\boldsymbol{X}_t, \boldsymbol{W}_t), \widetilde{f}_t(\boldsymbol{W}_t, \widehat{\boldsymbol{W}}_t)\Big)\,. \tag{4.45}$$

Of course, more realism pushes for incorporating more delays — $\boldsymbol{W}_{t+1} = \widetilde{f}_t(\boldsymbol{W}_t, \dots, \boldsymbol{W}_{t-k}, \widehat{\boldsymbol{W}}_t)$ — but at the price of increasing the dimension of the information state, now becoming the large vector $\big(\boldsymbol{X}_t, \boldsymbol{W}_t, \dots, \boldsymbol{W}_{t-k}\big)$, hitting the wall of the curse of dimensionality.

If the problem is written on a tree, we can write DPE with the couple physical state x and current node (identified with past noises).

Some approaches mix DP and a state of rather large dimension. For instance, *Stochastic Dual Dynamic Programming Algorithm* (SDDP) makes assumption on the objective function J (convexity) and on the dynamics functions F_t (linearity). With these, the value functions are shown to be convex, so that they can be approximated from below by the class of suprema of finite sets of linear functions. Such a structural property is a mean to partially overcome the curse of dimensionality of DP. In §4.3, we present SDDP as a DP approach where information is encoded in a tree and where value functions are cleverly approximated. Instead of computing the value function for any possible value of the state, the SDDP algorithm iteratively forges approximations of the value function that are improved around the states visited by optimal trajectories.

4.1.5 Summary table

In Table 4.1, we gather the decompositions listed above. It happens that all the decomposition methods we looked at are parallel, except the Dynamic Programming approach (SDDP being a DP like approach). Indeed, chained decomposition is intimately related to the natural flow of stages. The parallel decompositions that we presented have been deduced from a price decomposition scheme for different constraints.

Interestingly, decompositions can be weaved together or mixed, opening the way for a large variety of methods. For instance, we will present and dissect in §5.8 the *Dual Approximate Dynamic Programming* method (DADP). With the distinctions we established between decompositions, DADP can be seen as a spatial decomposition, where subproblems can be solved by time decomposition. More precisely, DADP makes it possible to solve subproblems by DP, rendering space and time decompositions compatible.

4.2 Progressive Hedging (PH)

We now briefly work out the Progressive Hedging Algorithm [56], discussed in §4.1.3, on the toy problem as depicted at end of §3.1.1. We assume that the demand D is a random variable and we minimize

$$\min_{q_0, q_1} \mathbb{E}[c_0 q_0 + c_1 q_1] \tag{4.46a}$$

	Decomposition			
	Time		Scenario	Space
	chained	parallel	parallel	parallel
Dynamic Programming	✓			
SDDP	✓			
DADP				✓
Progressive Hedging			✓	
Stochastic Pontryaguin			✓	

Table 4.1: Decomposition Methods

under the constraints

$$\begin{array}{ll} 0 & \leq q_0 \leq q_0^\sharp \\ 0 & \leq q_1 \\ D & \leq q_0 + q_1 \\ q_1 & \text{depends upon } D\ . \end{array} \tag{4.46b}$$

Since we have assumed that the decision q_1 depends upon the random variable D it follows that the optimal value, when minimizing with respect to q_1 for a fixed value of q_0, is

$$q_1^\sharp = [D - q_0]_+\ . \tag{4.47}$$

The minimization problem is thus equivalent to the following optimization problem

$$\min_{q_0} \mathbb{E}[c_0 q_0 + c_1 [D - q_0]_+] \tag{4.48a}$$

under the constraints

$$0 \quad \leq q_0 \leq q_0^\sharp \tag{4.48b}$$

Moreover, in order to simplify the exposition, we assume that the demand is a discrete valued random variable ($\mathbb{P}\{D = D_i\} = \pi_i$ for $i \in \{1, \cdots, N\}$), to obtain a discrete optimization problem

$$\min_{q_0} \sum_{i=1}^{N} \pi_i \left(c_0 q_0 + c_1 [D_i - q_0]_+\right) \tag{4.49a}$$

under the constraints

$$0 \quad \leq q_0 \leq q_0^\sharp \tag{4.49b}$$

If the value q_0 was allowed to change for each possible value of the random variable D, the problem could be easily solved. Here, we have a set of N potential problems which are coupled by the fact that they must use the same control q_0. Note that this coupling constraint is the so-called non-anticipative constraints.

The idea behind the Progressive Hedging algorithm is to break the coupling by dualization in order to decompose the original problem into N optimization problems. We thus introduce

N new variables $(q_{0,i})_{i\in\{1,\cdots,N\}}$ which are constrained to be all equal to the same value q_0. The optimization problem can be rewritten as follows

$$\min_{q_0,q_{0,i}\text{for all}i\in\{1,\cdots,N\}} \sum_{i=1}^{N} \pi_i \left(c_0 q_{0,i} + c_1[D_i - q_{0,i}]_+\right) \tag{4.50a}$$

under the constraints

$$\begin{array}{l} 0 \le q_0 \le q_0^\sharp \\ 0 \le q_{0,i} \le q_0^\sharp \quad \text{for all } i \in \{1,\cdots,N\} \\ (q_{0,i} - q_0) = 0 \quad \text{for all } i \in \{1,\cdots,N\} \end{array} \tag{4.50b}$$

where a new set of constraints has been introduced, and where q_0 has been replaced by $q_{0,i}$ in the cost, which is justified since $q_{0,i}$ is equal to q_0 for all $i \in \{1,\cdots,N\}$.

Using an augmented Lagrangian method, we dualize the set of coupling constraints. The resulting dual problem becomes

$$\max_{p_1,\cdots,p_N} \mathcal{D}(p) \tag{4.51a}$$

where $\mathcal{D}(p)$ is defined as follows

$$\mathcal{D}(p) = \min_{q_0,(q_{0,i})_{i\in\{1,\cdots,N\}}} \sum_{i=1}^{N} \pi_i \left(c_0 q_{0,i} + c_1[D_i - q_{0,i}]_+ + p_i(q_{0,i} - q_0) + \frac{r}{2}\|q_{0,i} - q_0\|^2\right) \tag{4.51b}$$

under the constraints

$$\begin{array}{l} 0 \le q_0 \le q_0^\sharp \\ 0 \le q_{0,i} \le q_0^\sharp \quad \text{for all } i \in \{1,\cdots,N\} \end{array} . \tag{4.51c}$$

The vector $p = (p_1,\cdots,p_N)$ is the vector of dual variables for the N coupling constraints, and $r > 0$ is a penalty parameter.

The algorithmic approach consists in maximizing the dual function (see the dual gradient algorithm in Remark 4.1) iteratively applying a gradient step:

$$p_i^{k+1} = p_i^k + r(q_{0,i}^{k+1} - q_0^{k+1}) \text{ for } i \in \{1,\cdots,N\}\,, \tag{4.52}$$

where $(q_0^{k+1}, q_{0,1}^{k+1}, \cdots, q_{0,N}^{k+1})$ are the minimizers in the problem $\mathcal{D}(p^k)$ and where p^0 is and arbitrary initial price vector.

A difficulty remains in the non-separability of the augmented Lagrangian due to the coupling variables q_0 appearing in the quadratic penalty term. The progressive hedging approach achieves decomposition by alternating minimization over the variable q_0 and over the set of variables $(q_{0,i})_{i\in\{1,\cdots,N\}}$. When q_0 is fixed to q_0^k, we compute $(q_{0,1}^{k+1},\cdots,q_{0,N}^{k+1})$ as a minimizer of $\mathcal{D}(p^k)$. This minimization splits into N independent optimization problems, which can be numerically solved using any adapted deterministic algorithm.

We then update q_0 by computing q_0^{k+1} as a minimizer of $\mathcal{D}(p^k)$, when $(q_{0,1},\cdots,q_{0,N})$ are fixed and set to $(q_{0,1}^{k+1},\cdots,q_{0,N}^{k+1})$. This last optimization problem can be solved explicitly and gives

$$q_0^{k+1} = \min\left(q_0^\sharp, \max\left(0, \sum_{i=1}^{N} \pi_i q_{0,i}^{k+1} + (\sum_{i=1}^{N} \pi_i p_i^k)/r\right)\right)\,. \tag{4.53}$$

Note that, when the initial price vector is such that $\sum_{i=1}^{N} \pi_i p_i^0 = 0$, then, using Equation (4.52), this property is maintained at any iteration. Therefore, the update for q_0 reduces to

$$q_0^{k+1} = \sum_{i=1}^{N} \pi_i q_{0,i}^{k+1} \ . \tag{4.54}$$

4.3 Stochastic Dual Dynamic Programming (SDDP)

The Stochastic Dual Dynamic Programming (SDDP) algorithm relies on Dynamic Programming. However, instead of computing the Bellman function at every state point, the algorithm constructs a polyhedral approximation of the Bellman function. More precisely, SDDP builds, for each time t, a piecewise linear function that is a lower approximation of the exact Bellman function.

The main asset of this approach consists in relying on linear programming solvers, that are numerically very efficient.

4.3.1 Preliminary results

We begin with a few preliminary results.

Piecewise linear functions as solutions of a linear program

Recall that the supremum of a finite number of affine functions is a convex piecewise linear function. We consider the function

$$\left\{ \begin{array}{rcl} g & : & \mathbb{R}^n \to \mathbb{R}, \\ g(x) & = & \max\limits_{i=1\dots k} \left(a_i \cdot x + b_i \right), \end{array} \right. \tag{4.55a}$$

where $b_i \in \mathbb{R}$ and $a_i \in \mathbb{R}^n$ for $i \in \{1, \cdots, k\}$. Then, we easily represent g as the solution of the linear program

$$\begin{array}{rcl} g(x) = \min\limits_{\alpha \in \mathbb{R}} & \alpha & \\ s.t. & \alpha \geq a_i \cdot x + b_i \ , & \forall i \in \{1, \cdots, k\} \ . \end{array} \tag{4.55b}$$

The cutting plane method: Kelley algorithm

The Kelley algorithm enables to approximate the minimum of a convex function φ on a compact C, with only an oracle able to compute, at each point, its value and derivative. Recalling that a convex function always lies above its tangents, the cutting plane method consists in constructing a *lower approximation* $\check{\varphi}$ of the objective function φ as the maximum of tangents. This lower approximation is a piecewise linear (affine) function, and is improved throughout the algorithm.

We now describe more precisely the Kelley algorithm. We start at a given point x_0. The first lower approximation is the tangent at this point,

$$\check{\varphi}^1(x_0) = T_1(x_0) \ . \tag{4.56a}$$

Then, at each step, we compute the equation of the tangent of φ at the minimum of the approximate function $\check{\varphi}$, as in

$$\left\{ \begin{array}{rcl} \underline{x}_k & \in & \arg\min_{x \in C} \check{\varphi}^k(x) \ , \\ T_k(x) & = & \varphi'(\underline{x}_k) \cdot (x - \underline{x}_k) + \varphi(\underline{x}_k) \ . \end{array} \right. \tag{4.56b}$$

The new lower approximation $\check{\varphi}^{k+1}$ is given by the maximum of the previous one $\check{\varphi}^k$ and the tangent T_k:

$$\check{\varphi}^{k+1}(x) = \max\left(\check{\varphi}^k(x), T_k(x)\right) \ . \tag{4.56c}$$

Finally, the approximate minimum is the minimum of the approximate function $\check{\varphi}^k$ at the end of the algorithm loop.

4.3.2 SDDP in the deterministic case

In a deterministic setting, we present the SDDP algorithm as an application of the Kelley's algorithm on a multistage problem, instead of a static problem.

Problem statement

We consider a controled dynamic system in discrete time that follows the equation

$$x_{t+1} = F_t(x_t, u_t) \ , \ \ \forall t \in \{0, \dots, T-1\} \ , \tag{4.57}$$

where $x_t \in \mathbb{X}$ is the *state* at time t of the system and $u_t \in \mathbb{U}$ the *control* applied at time t; F_t is the *dynamics* of the system. We assume that the sets $\mathbb{U}$ and $\mathbb{X}$ are compact subsets of a finite dimensional vector space. We also assume that the functions F_t are linear. At each time t, there is a convex *cost* $L_t(x_t, u_t)$ that depends on the current state and control. Moreover, there is a final cost $K(x_T)$ that depends on the final state of the system. A *policy* is a sequence of functions $\pi = (\pi_1, \dots, \pi_{T-1})$ giving, for any state x, a control u.

We consider the following problem

$$\begin{array}{rll} & \min\limits_{u \in \mathbb{U}^T} & \sum\limits_{t=0}^{T-1} L_t(x_t, u_t) + K(x_T), \\ s.t. & & x_{t+1} = F_t(x_t, u_t) \ . \end{array} \tag{4.58}$$

This problem can be solved by dynamic programming. For this purpose, we introduce the Bellman function defined by

$$\left\{ \begin{array}{rcl} V_T(x) & = & K(x), \\ V_t(x) & = & \min\limits_{u_t \in \mathbb{U}} L_t(x, u_t) + V_{t+1} \circ F_t(x, u_t) \ , \end{array} \right. \tag{4.59}$$

and an optimal policy is given by

$$\pi_t(x) \in \arg\min_{u_t \in \mathbb{U}} L_t(x, u_t) + V_{t+1} \circ F_t(x, u_t) \ , \ \ \forall t = 0, \dots, T-1 \ . \tag{4.60}$$

Bellman operator

Now, we find it useful to introduce the *Bellman operator* $\mathcal{T}_t$ as follows. For any time t, and any function A mapping the set of states into $\mathbb{R}$, we define:

$$\forall x \in \mathbb{X} \ , \ \ \mathcal{T}_t(A)(x) = \min_{u_t \in \mathbb{U}} \quad L_t(x, u_t) + A \circ F_t(x, u_t) \ . \tag{4.61}$$

Thus, the Bellman equation (4.59) simply reads

$$\left\{ \begin{array}{rcl} V_T & = & K, \\ V_t & = & \mathcal{T}_t(V_{t+1}) \ . \end{array} \right. \tag{4.62}$$

The Bellman operator displays the following properties.

Monotonicity. For any couple of functions $(V, \overline{V})$, we have

$$\forall x \in \mathbb{X} \ , \ \ V(x) \leq \overline{V}(x) \quad \Rightarrow \quad \forall x \in \mathbb{X}, \quad \big(\mathcal{T}V\big)(x) \leq \big(\mathcal{T}\overline{V}\big)(x). \tag{4.63}$$

Convexity. For any function V, if L_t is jointly convex in (x, u), if V is convex, and if F_t is affine, then

$$x \mapsto \big(\mathcal{T}V\big)(x) \quad \text{is convex.} \tag{4.64}$$

Linearity. For any piecewise linear function V, if L_t is jointly convex in (x, u) and is piecewise linear, and if F_t is affine, then

$$x \mapsto \big(\mathcal{T}V\big)(x) \quad \text{is piecewise linear.} \tag{4.65}$$

Brief recall on duality

Let J be a function mapping $\mathbb{X} \times \mathbb{U}$ into $\mathbb{R}$, supposed to be jointly convex in (x, u). Consider the function φ defined as

$$\varphi(x) = \min_{u \in \mathbb{U}} J(x, u) \ . \tag{4.66}$$

Then, the interpretation of the multiplier as a marginal cost implies that a subgradient $\lambda \in \partial\varphi(x_0)$ of φ at x_0 is the dual multiplier of the constraint $x_0 - x = 0$ in the following problem

$$\begin{array}{rll} & \min\limits_{x,u} & J(x, u) \ , \\ s.t. & & x_0 - x = 0 \ . \end{array} \tag{4.67}$$

In particular, it means that, as φ is convex,

$$\varphi(x) \geq \varphi(x_0) + \langle \lambda, x - x_0 \rangle \ , \ \ \forall x \in \mathbb{X} \ . \tag{4.68}$$

SDDP algorithm (deterministic case)

As in the Kelley's algorithm, the SDDP algorithm recursively constructs a lower-approximation $\check{V}_t^{(k)}$ of each Bellman function V_t as the supremum of a number of affine functions. Each of these affine functions is called a *cut*.

Stage k of the SDDP algorithm goes as follows. Suppose that we have a collection of T lower approximations of the Bellman functions, denoted $\check{V}_t^{(k)}$, each being the maximum of affine functions. Then, to improve our approximation of the Bellman function, we follow an optimal trajectory $t \mapsto x_t^{(k)}$ of the approximated problem, and we add a cut computed along the optimal trajectory for each Bellman function.

Thus, we begin with a loop forward in time, by setting $t = 0$ and $x_t^{(k)} = x_0$, and solve

$$\begin{array}{ll} \min\limits_{x,u} & L_t(x,u) + \check{V}_{t+1}^{(k)} \circ F_t(x,u) \ , \\ & x = x_t^{(k)} \ . \end{array} \tag{4.69}$$

We call $\beta_t^{(k+1)}$ the value of the problem, $\lambda_t^{(k+1)}$ a multiplier of the constraint $x = x_t^{(k)}$, and $u_t^{(k)}$ an optimal solution, so that we have

$$\left\{ \begin{array}{ll} \beta_t^{(k+1)} & = \mathcal{T}_t\left(\check{V}_{t+1}^{(k)}\right)\left(x_t^{(k)}\right) \ , \\ \lambda_t^{(k+1)} & \in \partial\mathcal{T}_t\left(\check{V}_{t+1}^{(k)}\right)\left(x_t^{(k)}\right) \ . \end{array} \right. \tag{4.70}$$

By (4.68), and by monotonicity of $\mathcal{T}_t$, we obtain

$$\beta_t^{(k+1)} + \langle \lambda_t^{(k+1)}, x - x_t^{(k)} \rangle \leq \mathcal{T}_t\left(\check{V}_{t+1}^{(k)}\right)(x) \leq \mathcal{T}_t\left(V_{t+1}\right)(x) = V_t(x) \ , \quad \forall x \in \mathbb{X} \ . \tag{4.71}$$

This ensures that the function $x \mapsto \beta_t^{(k+1)} + \left\langle \lambda_t^{(k+1)}, x - x_t^{(k)} \right\rangle$ is indeed a cut, i.e an affine function below V_t. Consequently, we update our approximation of V_t by defining

$$\check{V}_t^{(k+1)} = \max\left\{ \check{V}_t^{(k)}, \beta_t^{(k+1)} + \left\langle \lambda_t^{(k+1)}, \cdot - x_t^{(k)} \right\rangle \right\} \ . \tag{4.72}$$

Furthermore, we can see that $\check{V}_t^{(k+1)}$ is convex and lower than V_t. We go to the next time step by setting

$$x_{t+1}^{(k)} = F_t\left(x_t^{(k)}, u_t^{(k)}\right) \ . \tag{4.73}$$

Upon reaching time $t = T$, we have completed iteration k of the algorithm.

Initialization and stopping rule. To initialize the algorithm, it would seem that we need a lower bound for all value functions. According to our assumptions, such bound always exists. However, it is not necessary to implement the algorithm. Indeed, we can choose $V_t^{(0)} = 0$ to compute the trajectories, and simply set $V_t^{(1)}$ equal to the first cut, which means that we "forget" $V^{(0)}$ in the maximum that determines $V_t^{(1)}$.

Note that, at any step k of the algorithm, we have a (non optimal) solution $(u^{(k)})_t$ of Problem (4.58), and we can estimate the quality of the solution. Indeed $V_0^{(k)}(x_0)$ is a lower bound of the optimal value of Problem (4.58). Moreover the simulated cost following the solution given by

$$\sum_{t=0}^{T-1} L_t\left(x_t^{(k)}, u_t^{(k)}\right) + K\left(x_T^{(k)}\right) \ ,$$

is an upper bound of the value of Problem (4.58). A reasonable stopping rule for the algorithm is given by checking that the (relative) difference of the upper and lower bound is smaller than some constant as they are both asymptotically exact bounds, i.e. both converge toward the value of Problem (4.58).

4.3.3 The SDDP algorithm (DOASA implementation)

Now, we present the Stochastic Dual Dynamic Programming (SDDP) algorithm — more precisely, the so-called DOASA implementation. The first implementation of SDDP goes back to [43]. After a few tentatives, a proof of convergence in the linear case was given in [45] and, in the convex case, by [28].

Now we introduce random variables in our problem, complexifying the algorithm in different ways:

- we need probabilistic assumptions;
- at each stage k, we need to do a forward phase that yields a trajectory $t \mapsto x_t^{(k)}$, and a backward phase that gives a new cut;
- we cannot compute an exact upper bound for the problem's value.

Problem statement

As in the deterministic case, we consider a controled stochastic dynamic system in discrete time that follows the equation

$$\boldsymbol{X}_{t+1} = F_t(\boldsymbol{X}_t, \boldsymbol{U}_t, \boldsymbol{W}_t) \ , \quad \forall t \in \{0, \dots, T-1\} \ , \tag{4.74}$$

where $\boldsymbol{X}_t \in \mathbb{X}$ is the *state* at time t of the system, $\boldsymbol{U}_t$ the *control* applied at time t, and $\boldsymbol{W}_t$ an exogeneous discrete *noise*; F_t is the *dynamic* of the system.

We assume that the stochastic process $(\boldsymbol{W}_t)_{t=1,\dots,T-1}$ is independent in time. Moreover, we assume that, at any time t, we know the whole past of the noise process up to, and including, time t. The independence assumption ensures that we can only consider control $\boldsymbol{U}_t$ measurable with respect to $(\boldsymbol{X}_t, \boldsymbol{W}_t)$. This is the so-called hazard-decision information structure, where we know at time t the realization of the noise before choosing our control.

At each time t, there is a jointly convex in (x_t, u_t) *cost* $L_t(x_t, u_t, W_t)$ that depends on the current state and control, and there is a final cost $K(x_T)$ that depends on the final state

of the system. A *policy* is a sequence of measurable functions $\pi = (\pi_1, \dots, \pi_{T-1})$ giving, for any state x and realization of the noise w, a control u.

We consider the following problem

$$\begin{array}{rll} \min\limits_{\pi} & \mathbb{E}\left(\sum\limits_{t=0}^{T-1} L_t(\boldsymbol{X}_t, \boldsymbol{U}_t, \boldsymbol{W}_t) + K(\boldsymbol{X}_T)\right) , \\ s.t. & \boldsymbol{X}_{t+1} = F_t(\boldsymbol{X}_t, \boldsymbol{U}_t, \boldsymbol{W}_t) , \\ & \boldsymbol{U}_t = \pi_t(\boldsymbol{X}_t, \boldsymbol{W}_t) . \end{array} \tag{4.75}$$

This problem can be solved by dynamic programming. For this purpose, we introduce the Bellman function defined by

$$\left\{ \begin{array}{lcl} V_T(x) & = & K(x) , \\ \widehat{V}_t(x, w) & = & \min\limits_{u_t \in \mathbb{U}} L_t(x, u_t, w) + V_{t+1} \circ F_t(x, u_t, w) , \\ V_t(x) & = & \mathbb{E}\left(\widehat{V}_t(x, \boldsymbol{W}_t)\right) . \end{array} \right. \tag{4.76}$$

Indeed, Dynamic Programming ensures that an optimal policy for this problem is given by

$$\pi_t(x, w) \in \arg\min_{u_t \in \mathbb{U}} L_t(x, u_t, w) + V_{t+1} \circ F_t(x, u_t, w) , \quad \forall t = 0, \dots, T-1 . \tag{4.77}$$

Bellman operator

As in the deterministic case, we also introduce the stochastic *Bellman operator* $\mathcal{T}_t$ of the problem. For any time t, and any function A mapping the set of states and noises $\mathbb{X} \times \mathbb{W}$ into $\mathbb{R}$, we define:

$$\widehat{\mathcal{T}}_t(A)(x, w) = \min_{u_t \in \mathbb{U}} L_t(x, u_t, w) + A \circ F_t(x, u_t, w) , \quad \forall x \in \mathbb{X} . \tag{4.78a}$$

The stochastic Bellman operator $\mathcal{T}_t$, is given by

$$\mathcal{T}_t(A)(x) = \mathbb{E}\big(\widehat{\mathcal{T}}_t(A)(x, \boldsymbol{W}_t)\big) , \quad \forall x \in \mathbb{X} . \tag{4.78b}$$

Thus, the Bellman equation 4.76 simply reads

$$\left\{ \begin{array}{lcl} V_T(x) & = & K(x) , \\ V_t(x) & = & \mathcal{T}_t(V_{t+1})(x) . \end{array} \right. \tag{4.79}$$

Duality theory

Suppose that, at iteration k and time step $t+1$, we have a lower estimation $\check{V}_{t+1}^{(k+1)}$ of V_{t+1}. We consider, for any w, the problem

$$\begin{array}{ll} \min\limits_{x,u} & L_t(x, u, w) + \check{V}_{t+1}^{(k+1)} \circ F_t(x, u, w) , \\ s.t & x = x_t^{(k)} , \end{array} \tag{4.80}$$

and call $\widehat{\beta}_t^{(k+1)}(w)$ the optimal value, and $\widehat{\lambda}_t^{(k+1)}(w)$ an optimal multiplier of the constraint. As in the deterministic case, we have

$$\left\{ \begin{array}{ll} \widehat{\beta}_t^{(k+1)}(w) & = \widehat{\mathcal{T}}_t \left(\check{V}_{t+1}^{(k)} \right) (x_t^{(k)}, w) \ , \\ \widehat{\lambda}_t^{(k+1)}(w) & \in \partial_x \widehat{\mathcal{T}}_t \left(\check{V}_{t+1}^{(k)} \right) (x_t^{(k)}, w) \ , \end{array} \right. \tag{4.81}$$

which gives

$$\widehat{\beta}_t^{(k+1)}(w) + \left\langle \widehat{\lambda}_t^{(k+1)}(w), x - x_t^{(k)} \right\rangle \leq \widehat{\mathcal{T}}_t \left(\check{V}_{t+1}^{(k)} \right) (x, w) \leq \widehat{V}_t(x, w) \ , \ \ \forall x \in \mathbb{X} \ , \ \ \forall \omega \ . \tag{4.82}$$

Thus, we obtain an affine minorant for each realization of $\boldsymbol{W}_t$. Replacing w by the random variable $\boldsymbol{W}_t$, and taking the expectation, yields the following affine minorant

$$\beta_t^{(k+1)} + \left\langle \lambda_t^{(k+1)}, \cdot - x_t^{(k)} \right\rangle \leq V_t \ , \tag{4.83a}$$

where

$$\beta_t^{(k+1)} = \mathbb{E} \left(\beta_t^{(k+1)}(\boldsymbol{W}_t) \right) \ , \tag{4.83b}$$

$$\lambda_t^{(k+1)} = \mathbb{E} \left(\lambda_t^{(k+1)}(\boldsymbol{W}_t) \right) \ , \tag{4.83c}$$

or, in other words,

$$\left\{ \begin{array}{ll} \beta_t^{(k+1)} & = \mathcal{T}_t \left(\check{V}_{t+1}^{(k)} \right) (x, w) \ , \\ \lambda_t^{(k+1)} & \in \partial_x \mathcal{T}_t \left(\check{V}_{t+1}^{(k)} \right) (x, w) \ . \end{array} \right. \tag{4.83d}$$

SDDP algorithm (stochastic case)

At the beginning of step k, we suppose that we have, for each time step t, an approximation $\check{V}_t^{(k)}$ of V_t satisfying

- $\check{V}_t^{(k)} \leq V_t$,
- $\check{V}_T^k = K$,
- $\check{V}_t^{(k)}$ is convex.

Forward path. We randomly select a scenario $(w_0, \ldots, w_{T-1}) \in \mathbb{W}^T$. We recursively, by a forward loop in time, define a trajectory $(x_t^{(k)})_{t=0,\ldots,T}$ by $x_{t+1}^{(k)} = F_t(x_t^{(k)}, u_t^{(k)}, w_t)$, where $u_t^{(k)}$ is an optimal solution of

$$\min_{u \in \mathbb{U}} L_t \left(x_t^{(k)}, u, w_t \right) + \check{V}_{t+1}^{(k)} \circ F_t \left(x_t^{(k)}, u, w_t \right) \ . \tag{4.84}$$

This trajectory is a trajectory given by the optimal policy of Problem (4.75) where V_t is replaced by its current approximation $\check{V}_t^{(k)}$.

Backward path. We improve the approximation of V_t by adding a cut at $V_t(x_t^{(k)})$. We procede backward in time by setting $t = T-1$, and solve Problem (4.80) for every w in the support of $\boldsymbol{W}_t$. We obtain $\widehat{\lambda}_t^{(k+1)}(w)$ and $\widehat{\beta}_t^{(k+1)}(w)$. As $\boldsymbol{W}_t$ is discrete, we can compute exactly the expectations $\lambda_t^{(k+1)}$ and $\beta_t^{(k+1)}$, and define the new lower approximation

$$\check{V}_t^{(k+1)}(x) = \max\left\{\check{V}_t^{(k)}(x), \beta_t^{(k+1)} + \left\langle \lambda_t^{(k+1)}, x - x_t^{(k)} \right\rangle\right\} \ , \ \forall x \in \mathbb{X} \, . \tag{4.85}$$

We can easily check that $\check{V}_t^{(k+1)}$ is a lower convex approximation of V_t, and go one step back in time: $t \leftarrow t-1$. Upon reaching $t = 0$, we have completed step k of the algorithm.

Initialization and stopping rule

We can initialize the algorithm as suggested in the deterministic case. However, to accelerate the convergence, it can be useful to bypass a few forward paths by arbitrarily choosing some trajectories $t \mapsto x_t^{(k)}$.

As in the deterministic case, we have an exact lower bound of Problem (4.75) given by $V_0^{(k)}(x_0)$. However, determining an upper bound is more difficult. Indeed, the upper bound is given as the expectation of the cost induced by the optimal strategy of the approximated problem. This expectation is to be taken over the whole process $(W_0, \ldots, W_{T-1})$ and cannot be computed exactly. Consequently, this expectation is estimated by Monte-Carlo methods — providing a confidence interval that is a random interval $[\hat{m}_n^{(k)} - e_n, \hat{m}_n^{(k)} + e_n]$ such that the actual expectation has 95% (for example) chance of being in. Note that the error e_n decrease with the number n of scenarios used in the Monte Carlo method.

A stopping rule can be defined in the following way. Choose an a priori error $\varepsilon > 0$ and stop the algorithm when the upper bound $\hat{m}_n^{(k)} + e_n$ of the confidence interval of the upper bound of the value of Problem (4.75) is less than ε above the (exact) lower bound $V_0^{(k)}(x_0)$. In this case, we have, say, 95% chance that the approximate policy given by the SDDP algorithm yields a value less than ε above the optimal value.

Other approaches

In the algorithm presented here, also called DOASA, we select one scenario (one realization of $(\boldsymbol{W}_1, \ldots, \boldsymbol{W}_{T-1})$) to do a forward and backward path. I t is also possible to select a number N of scenarios to do the forward path (computations can be parallelized). Then, during the backward path, we add N cuts to V_t before computing the cuts on V_{t-1}. This is the original implementation of the algorithm.

Otherwise, the CUPPS algorithm suggests to use $\check{V}_{t+1}^{(k)}$ instead of $V_{t+1}^{(k+1)}$ in the computation of the cuts (4.81). In practice, it means that we do only a forward path at step k:

- select randomly a scenario $(w_t)_{t=0,\ldots,T-1}$;
- at time t, we have a state $x_t^{(k)}$, and we solve problem (4.80) for every possible realization of W_t to compute the new cut for V_t;

- choose the optimal control corresponding to the realization $W_t = w_t$ in order to compute the state $x_{t+1}^{(k)}$ where the cut for V_{t+1} will be computed, and go to the next step.

Moreover, it is possible to select more than one scenario, and everything can be parallelized in this implementation.

Finally it is always possible, and most of the time fruitful to compute some cuts before starting the algorithm. One way of doing this is to bypass the forward phase by choosing the trajectory $(x_t^{(k)})_{t=0,\dots,T}$ instead of computing it as some optimal trajectory.

4.3.4 From data to application of SDDP: a number of approximations

We conclude this presentation of the SDDP algorithm by quickly showing how the algorithm can be applied in practice. More information and numerical results can be found in a technical report — by A. Shapiro, W. Tekaya, J. P. da Costa and M. Pereira Soares — where SDDP is applied for the Brazilian electrical system operator [65].

Four layer of errors (or approximation) appear when using an SDDP approach: modeling error, discretization error, finite running time error and Monte-Carlo error. Let us review them.

From historical data and through the use of classical statistical tools, an ARMA model over the random noises of the system is fitted (ARMA because of the linearity assumptions in the dynamics). It is impossible to quantify the modeling error between the "real random process" and the ARMA model. We turn and discretize this continuous distribution into a huge scenario tree. This discretization corresponds to a *Sampling Average Approximation* approach, and the error between the ARMA model and its quantified version can be controlled. The algorithm will only run for a finite number of steps, hence there is a gap between the policy found and the optimal policy. Finally, when a policy is evaluated, the exact expectation cannot be computed and is estimated by a Monte Carlo approach. This last error can also be controled.

4.4 Approximate Dynamic Programming

As we have seen in §4.1.4, a multistage optimization problem can theoretically be solved by a Dynamic Programming approach. Unfortunately, the practical use of Dynamic Programming is limited by the size of the state of the problem (the so-called *curse of dimensionality*). In a smart-grid management problem, the state of the system is generally too big for numerical resolution of the system (think of having electrical storage in different nodes of a network). One way of dealing with this problem consists in approximating the Bellman function by a function with a small number of arguments. The field of *Approximate Dynamic Programming* is working out such approach of large scale problems. We refer the reader to [52] for a short introduction to this field and to the outline hereafter.

We review a few of the mains ideas of ADP. The first idea consist in trying to have approximation of the value functions instead of exact value function, and improving them

step after step. Note that the SDDP algorithm has the same approach of the problem. As in SDDP, a given approximation of the value functions yields a strategy. Hence, the set of all strategies considered is related to the set of all possible approximated value function considered. The second idea consist in going forward in time instead of backward (as in the classical dynamic programming approach). Indeed, when going backward, the value function is exactly computed for every possible state, even those states that the system will not visit. On the other hand, a forward approach allows to focus the numerical effort on the states that are actually visited by the system. Another idea consists in focusing on the so-called post-decision state, that is, the state achieved after a given decision.

Central to ADP is the use of the approximate value function for making decisions. Finding a good strategy to obtain an approximation remains an art. We gives some general ideas that work for the broadest possible classes. The first strategy consists in aggregating states. Roughly speaking, we consider a partition of the set of states. When we visit (in a forward approach) a given state and can update the value function at this point, we update in the same way the value function of all states in the same class of the partition. A second generic strategy consists in choosing a basis of functions of the state and stating that the approximate value functions are linear combination of these basis elements. Then, the challenge consists in fitting the coefficients of the linear combination of the basis. Classical approaches for this rely on stochastic gradient algorithm, recursive least square or Kalman filter.

Finally, note that one of the downfalls of the ADP approach is that we have no idea of how far from optimum is the solution obtained. To evaluate an ADP-obtained policy, we can either compare it to the optimal one on a toy model, or to an optimal deterministic or myopic strategy. Moreover, ADP can require fine tuning, even if there exists some generic rules to follow.

Chapter 5

Relating Ongoing Works and Practices

Here, we highlight current work and research by academic colleagues, at the international level. The examples below display a partial picture of the contemporary front tackling issues of interest for the energy sector when stochasticity is explicitely taken into account.

5.1 Stochastic Unit Commitment at ISO Scale – An ARPAe Project (Roger Wets and David Woodruff, UC Davis)

Project partners Alstom Grid, Iowa State University, ISO-New England, Sandia National Laboratories and the University of California Davis have embarked on an ARPAe (US Department of Energy, GENI project) funded project led by Sandia to develop methods for solving stochastic unit commitment (UC) problems at ISO scale including specification of scenarios for uncertain demand and high penetration wind power. Using formulations validated against Alstom's test models and data from ISO New England (and other sources for high wind penetration), we are estimating what energy savings would have occurred had ISO-NE been using stochastic UC during the test-year 2011, but now including renewables-supply, in particular, high wind penetration.

To ensure feasibility and adaptability, the computational platform has been designed to be flexible, modular, and extensible. The word "extensible" is taken to mean the ability to modify parts of the platform without affecting other parts, and to scale up the platform to larger sizes with relative ease. To maintain computational feasibility for practical application, the scope of the computational platform have been kept within reasonable; i.e., we are using hundreds, but not thousands of CPUs in parallel.

Unit commitments represent now-decisions (on day $D-1$), which must be determined before values of uncertain parameters such as available intermittent resources, variable generator output, unscheduled outages, and loads are known. Economic dispatch decisions (on

day D) provide "recourse" decisions when the resolutions of the (uncertain) parameters are at hand.

Scenario trees are used to represent the possible combinations of parameter values (scenarios), in conjunction with their probabilities of realization. Advanced scenario creation techniques have been leveraged to compute cost confidence intervals and to minimize the number of scenarios required to achieve reliable solutions.

Innovative decomposition strategies have been leveraged to meet run-time and memory requirements, by distributing sub-problem solutions across distinct computer nodes. Specifically, Progressive Hedging is a decomposition scheme that works well with mixed integer variables and specific non-linear structures, and is naturally parallelizable.

Great care is needed to manage scenario tree size, which directly impacts computational difficulty. Mitigation approaches include generating a parsimonious tree with a coarse state space, and reducing its size to approximately get the same solution with fewer scenarios.

The final report for the project is due during the early of Autumn of 2014.

5.2 Clearing the Jungle of Stochastic Unit Commitment (Warren B. Powell, Princeton)

With the growth in interest in wind and solar energy, there has been renewed interest in solving unit commitment problems in a way that handle the high level of uncertainty introduced by renewables. This has produced a burst of interest in a particular algorithmic strategy developed by the stochastic programming community which has become known (confusingly) as the "stochastic unit commitment problem" which requires solving very large-scale integer programming problems. Other algorithmic strategies with names such as SDDP (stochastic dual dynamic programming), and ADP (approximate dynamic programming) are also being tossed around as alternatives to what grid operators are doing today which are (incorrectly) characterized as "deterministic models."

In this contribution, we are going to make the following points:

- The "stochastic unit commitment problem", as it is widely known in the literature, is actually a policy based on solving an approximate lookahead model, which is a form of rolling horizon procedure (the same class as a deterministic lookahead model). The stochastic program is very large and hard to solve, yet an optimal solution to this problem is *not* an optimal policy.

- There is widespread confusion between what we call a *stochastic base model* which is typically implemented as a simulator (grid operators use the real world as their simulator), and a stochastic lookahead model, which is the basis of a particular class of policies derived using stochastic programming.

- There are four fundamental classes of policies that can be used to solve any stochastic optimization problem: PFAs (policy function approximations), CFAs (robust cost function approximations), VFAs (policies based on value function approximations),

and lookaheads (which solve some approximation of the future). These can be mixed to form hybrids. The proper way to evaluate any policy is through the base model (which is the stochastic version of the objective function for a deterministic model), which is typically implemented as a simulator.

- Stochastic programming (often referred to as the "stochastic unit commitment problem") depends on a sampled representation of the probability model, which prevents modelers from capturing important structural properties. Not only are these models very hard to solve, we now believe that this problem class is fundamentally ill-suited to the use of scenario-based methods (explained below).

- Stochastic dual dynamic programming (SDDP), and other methods that are characterized as approximate dynamic programming, are all policies based on value function approximations, which may work well for hydro storage problems (with issues), but are poorly suited to the stochastic unit commitment problem.

- Our research is leading us to the conclusion that a form of robust cost function approximation will prove to be the best path for stochastic unit commitment. Robust CFAs represent the class of policies already in widespread use by grid operators. These will need to be tuned, ideally in a high quality simulator (known as the stochastic base model), but we believe that this is correct approach.

Below, we support these conclusions with additional discussion.

5.2.1 Modeling stochastic optimization problems

There is widespread agreement about the modeling of deterministic optimization problems, but the situation becomes much more complex when random variables are introduce. At this point, the research community tends to fragment into fields with names like stochastic programming, dynamic programming, reinforcement learning, stochastic search, simulation optimization, and optimal control.

There are five fundamental elements to any sequential stochastic optimization problem. These are:

- State S_t - This represents what we know (more precisely, what we need to know) at time t before we make a decision (more on this below).

- Actions - Depending on the community, these might be discrete actions a, continuous controls u, or decisions x (which are typically vectors, and might be continuous, integer or a mixture). We determine a decision using a *policy*, which is a rule (or function) for making a decision. If our decision is action a_t, we designate the policy as the function $A^\pi_t(S_t)$ (or $A^\pi(S_t)$ if our policy is stationary). If we are using decision x_t, we use the function $X^\pi_t(S_t)$.

- Exogenous information W_t - Starting at time 0, we observe exogenous information (prices, demands, equipment breakdowns) as the sequence $W_1, W_2, \ldots$. We use the convention that any variable indexed by t is known at time t.

- The transition function - We write this as $S_{t+1} = S^M(S_t, x_t, W_{t+1})$, where $S^M(\cdot)$ has been referred to as the system model, plant model, transfer function or just model, but we refer to it as the transition function, consisting of the equations that describe the evolution of the system from t to $t+1$. The transition function may include (controllable) systems of linear equations, or the exogenous evolution of prices, weather, and customer demands.

- The objective function - We assume we are given a cost/reward/utility function that we refer to generically as the *contribution function* which may depend on the state S_t and the action x_t, so we write it as $C(S_t, x_t)$. The objective requires finding the policy that minimizes expected costs, which is written

$$\min_{\pi \in \Pi} \mathbb{E}^{\pi} \sum_{t=0}^{T} C(S_t, X_t^{\pi}(S_t)), \tag{5.1}$$

  where $S_{t+1} = S^M(S_t, x_t, W_{t+1})$, and where the expectation is over all possible sequences $W_1, W_2, \ldots, W_T$, which may depend on the actions taken. The goal here is to find the best policy, which means that we are looking for the best *function* for making a decision. There is growing interest in replacing the expectation with a quantile, risk measures or the worst outcome (robust optimization).

The most important equation in this model is equation (5.1), which represents the objective function for a problem where we are looking for the best *policy*. In contrast with deterministic optimization problems where we are looking for the best *decision* x, in (sequential) stochastic optimization problems we are looking for the best *function* (that is, policy) for determining a decision given the information in the state S_t.

Most of the arguments about how to solve stochastic unit commitment problems are basically arguments about which class of policy is best. We are going to first summarize what appear to be four fundamental classes of policies, and we then touch on the policies which are attracting the most attention.

5.2.2 The four classes of policies

We feel that the reason that so few authors write down the objective function (5.1) is that the problem of "searching over policies" looks vague and abstract, with no obvious path to computation.

Our experience is that every paper (with which we are familiar) solving a (sequential) stochastic optimization problem ends up using one of the four following classes of policies:

- Policy function approximations (PFAs) - A policy function approximation represents some analytic function that does not involve solving an optimization problem. A PFA is typically a simple function such as "pump water into the hydro facility at night, release during the middle of the day" or "store energy in the battery if the price is below p^{buy}, and sell it when it goes above p^{sell}. It might be a simple parametric model such as

$$X^{\pi}(S_t|\theta) = \theta_0 + \theta_1 S_t + \theta_2 S_t^2. \tag{5.2}$$

  A popular strategy in engineering is to use neural networks (see [77] and [68]).

- Optimizing a cost function approximation (CFAs) - There are problems where a simple myopic policy can produce good (in rare cases optimal) results. A myopic policy would be written

$$X_t^{\pi}(S_t|\theta) = \arg\min_{x\in\mathcal{X}_t} C(S_t, x).$$

  Not surprisingly, this would rarely work well in practice. However, there are problems where a slightly modified cost function might work quite well. One class of approximations looks like

$$X_t^{\pi}(S_t|\theta) = \arg\min_{x\in\mathcal{X}_t} \Big(C(S_t, x) + \sum_{f\in\mathcal{F}} \theta_f \phi_f(S_t, x)\Big). \tag{5.3}$$

  Here, $(\phi_f(S_t, x))_{f\in\mathcal{F}}$, is a set of *basis functions* (as they are known in the approximate dynamic programming community) which might be of the form $S_t x$, $S_t x^2$, x, x^2, which serves as a type of correction term. However, there are other problems where we make direct changes to the cost function itself, or perhaps the constraints (e.g., imposing a minimum inventory level). We can represent this class of policies more broadly by writing

$$X_t^{\pi}(S_t|\theta) = \arg\min_{x\in\mathcal{X}_t} \overline{C}^{\pi}(S_t, x|\theta). \tag{5.4}$$

  where $\overline{C}_t^{\pi}(S_t, x|\theta)$ is some sort of parametric approximation. Here, we would let π carry the information about the structure of the approximation (such as the basis functions in equation (5.3)) and we let θ capture all tunable parameters. A CFA can represent costs at a point in time, or over a horizon.

- Policies that depend on a value function approximation (VFAs) - These are the policies most often associated with dynamic programming, and are written

$$X_t^{\pi}(S_t|\theta) = \arg\min_{x\in\mathcal{X}_t} \Big(C(S_t, x) + \mathbb{E}\{\overline{V}_{t+1}(S_{t+1}|\theta)|S_t\}\Big). \tag{5.5}$$

  where $S_{t+1} = S^M(S_t, x_t, W_{t+1})$. Since expectations can be notoriously hard to compute (imagine if our random variable W_t has, say, 100 dimensions), we can use the device of

the post-decision state variable (the state immediately after a decision has been made, but before any new information has arrived). Let S_t^x be the post-decision state, which means it is a deterministic function of x_t. This allows us to write

$$X_t^\pi(S_t|\theta) = \arg\min_{x\in\mathcal{X}_t} \big(C(S_t,x) + \overline{V}_t(S_t^x|\theta)\big). \tag{5.6}$$

In both (5.5) and (5.6), we have to create an approximation of the value function. Again, we assume that the index π captures the structure of the function, as well as any tunable parameters represented by θ. For example, a popular approximation is linear regression. Assume that someone has devised a series of explanatory variables ("basis functions") $\phi_f(S_t^x)$ for $f \in \mathcal{F}$. Then we can write

$$X_t^\pi(S_t|\theta) = \arg\min_{x\in\mathcal{X}_t} \big(C(S_t,x) + \sum_{f\in\mathcal{F}} \theta_f\phi_f(S_t^x)\big). \tag{5.7}$$

- Lookahead policies - The simplest lookahead policy involves optimizing over a horizon H deterministically. Let $\widetilde{x}_{tt'}$ represent the decision variables (this might be a vector) for time t' in the lookahead model that has been triggered at time t. Variables with tildes represent the lookahead model, so we do not confuse them with the base model. A deterministic lookahead policy might be written

$$X_t^\pi(S_t|\theta) = \arg\min_{\widetilde{x}_{tt},\ldots,\widetilde{x}_{t,t+H}} \sum_{t'=t}^{t+H} C(\widetilde{S}_{tt'},\widetilde{x}_{tt'}). \tag{5.8}$$

Normally, we optimize over the horizon $(t,\ldots,t+H)$ but only implement the first decision, so $X_t^\pi(S_t|\theta) = \widetilde{x}_{tt}$. All variables in the lookahead model are indexed by (tt') where t represents when the decision is being made (and therefore its information content), while t' is when the decision impacts the physical system. The lookahead variables (with tildes) may capture various approximations; for example, our base model may step forward in 5-minute increments, while the lookahead model may uses hourly increments so it is easier to solve. Here, the parameter θ captures all the parameters that determine the formulation of the lookahead model (including choices such as the planning horizon).

We might also use a stochastic lookahead model

$$\min_{\widetilde{x}_{tt},(\widetilde{x}_{tt'}(\widetilde{\omega}),t<t'\le t+H),\forall\widetilde{\omega}\in\widetilde{\Omega}_t} \left(\widetilde{c}_{tt}\widetilde{x}_{tt} + \sum_{\widetilde{\omega}\in\widetilde{\Omega}_t} p(\widetilde{\omega}) \sum_{t'=t+1}^{t+H} \widetilde{c}_{tt'}(\widetilde{\omega})\widetilde{x}_{tt'}(\widetilde{\omega})\right). \tag{5.9}$$

In this case, θ captures parameters such as the number of information stages and the number of samples per stage (this community refers to the elements of $\widetilde{\Omega}_t$ as *scenarios*). An "information stage" consists of revealing information, followed by making decisions that use this information. We need to construct a lookahead stochastic process, captured by the set $\widetilde{\Omega}_t$, to differentiate stochastic *scenarios* in the lookahead model from

the sample path $\omega \in \Omega$ that we might be following in the base model (we suggest using "sample path" to describe the evolution of information in the base model, and "scenario" to represent the evolution of information in the lookahead model). The choice of the number of stages, and the construction of the set $\widetilde{\Omega}_t$, represent important decisions in the design of the lookahead model, which we parameterize by θ.

These four classes of policies now provide meaning to the "search over policies" in equation (5.1). We can test policies within each of the four classes, as well as combinations. For a given class C, we have to search over some parameter vector $\theta \in \Theta^C$ which governs the specific behavior of a policy.

At this point, we have to use some intuition about the nature of the problem to guide this search. For example, policy function approximations are bested suited for relatively low dimensional decisions (e.g., store energy in the battery when prices are below some level, sell when they are above a higher level).

Value function approximations tend to work well when we can reasonably approximate the value of being in a state S_{t+1} starting at time $t+1$ and moving forward. This has been attractive to people working on hydro storage problems (see, e.g., [43] and [66]), but these are problems where the state variable (in particular, the variable describing the physical state of the system) is relatively simple, making $V_{t+1}(S_{t+1})$ easy to approximate

The stochastic unit commitment problem, on the other hand, is a very complex problem from the perspective of the physical state. It involves modeling the state of each generator. Generators that are on need to stay on for a period of time; once a generator is turned off, it needs to stay off for a period of time. Generators needed to be ramped up or down at rates limited by the technology.

For this reason, the stochastic unit commitment problem has to be solved with some form of lookahead policy. Below we discuss two strategies: stochastic lookahead policies, and robost cost function approximations.

5.2.3 Stochastic lookahead policies

Many have assumed that the right approach to handle uncertainty in a unit commitment problem (which is a large integer program) should be a stochastic lookahead policy (such as that given in equation (5.9) (see [70], [71], [58], [58]). Here, we represent the future using a set of scenarios given by the set $\widetilde{\Omega}_t$, which might describe possible realizations of wind energy starting at time t, extending over some horizon to $t+H$. This formulation has attracted considerable attention from the research community, especially from the national laboratories in the U.S. because it is viewed as a potential application of their super computers.

In very recent research at PENSA Laboratory at Princeton University[1] we performed a very detailed analysis of the value of advance information in the context of a study of the value of off-shore wind. This work has shown that the value of perfect information is quite high; wind is very hard to forecast, even just an hour into the future. In one

[1] http://energysystems.princeton.edu

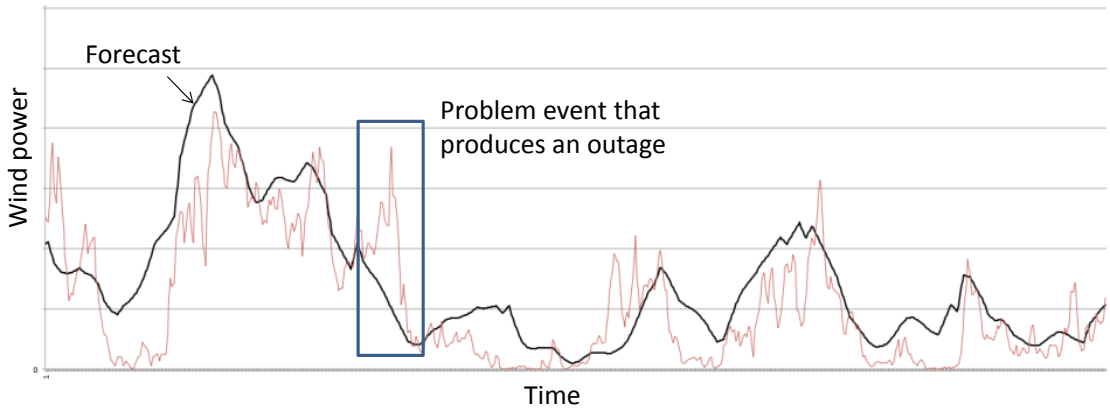


Figure 5.1: Sample path showing sharp rise and drop over the forecast.

example, perfect information would have told us that we should schedule extra steam in the day-ahead market, because we would otherwise run out of reserve gas turbines in the hour-ahead market. However, this is captured only when a scenario includes a specific behavior, illustrated in figure 5.1, which highlights a particular period where the wind rose above the forecast, and then dropped sharply. In our example, we could not schedule enough gas turbine reserves (access to a hydro facility would have been wonderful). We could solve the problem by scheduling more steam the day before, but with this scenario, our model would only schedule the steam at the particular point in time that the event happened.

In reality, we know that this event could happen at any point in time. If we depended on scenarios to create this behavior, we would need a very large set of scenarios, far beyond what even current research is considering for this problem class.

The real problem of scenario trees is that we are using scenarios to do a better job scheduling steam generation, which is a very high-dimensional vector. Most testing of so-called stochastic unit commitment models are being done with only a few dozen scenarios, but this is being used to determine the steam generation schedule. PJM Interconnections has 300 steam generators, and we have to make scheduling decisions each hour of the day, giving us 7,200 variables. You are not going to be able to come up with robust decisions on 7,200 variables using 20 scenarios.

5.2.4 Robust cost function approximations

Robust cost function approximations involve replacing the original cost function in a lookahead model, such as that given in equation (5.8), with a modified cost function that includes reserve generation. If $L_{tt'}$ is the forecast of the load at time t' (made at time t), we might insist that we have $\theta^{up}L_{tt'}$ of up-ramping reserves, and $\theta^{down}L_{tt'}$.

We refer to this as a *parametric cost function approximation* and let $X^{LA-det}(S_t|\theta)$ be the resulting solution given the control parameters $\theta = (\theta^{up}, \theta^{down})$. Our cost function is now parameterized by θ, and we now have to find the value of θ that best solves our objective

function (5.1). Now, the search over policies π has become a search over the parameters θ.

Solving the problem this way provides us with two important features: 1) The resulting optimization problem is comparable in complexity to the deterministic lookahead policy, so it is computationally tractable. 2) By imposing structure on the design of our cost function approximation, we get the behavior that we want to schedule extra reserve at *every* time of day, and not just when a scenario happened to represent the wind condition that causes us difficult.

Readers familiar with the unit commitment problem may quickly recognize that what we are calling a policy based on a robust cost function approximation is approximately what the grid operators are already doing! Does this mean that they can already handle high levels of renewables? The correct answer is a highly-qualified yes, but only up to a point.

In our experiments simulating planning at PJM Interconnections in the U.S., we have found that PJM's current policy cannot handle high levels of wind, but it can be tuned to handle much higher levels, but only up to a certain level, which is comparable to getting approximately 20 percent of their energy from wind. We believe that they can achieve higher levels with a slight modification to their robust cost function approximation, but at some point they need to include other strategies such as additional gas turbines, fast demand response, or large amounts of energy storage.

5.2.5 Parametric versus nonparametric approximations

When we compare the approach based on scenario trees to the approach based on robust cost function approximations, we find that there is a sharp contrast in the style of approximations. Scenario trees can be described as a *nonparametric* method of representing the underling set of outcomes. By contrast, the robust cost function approximation is a *parametric* approximation where a person has to choose the structure (as happens with all parametric models) while computers tune the parameters.

Our experience with grid operators is that they tune their parameters in an ad hoc manner, using the real world as their simulator. When testing high penetrations of wind, it is necessary to use a simulator, and we have spent years developing and calibrating SMART-ISO (see http://energysystems.princeton.edu/smartiso.htm).

5.2.6 Closing notes

Despite the considerable interest in solving the very large stochastic integer programs that arise when using stochastic programming, we believe that we can do a much better job using parametric cost function approximations, tuned using a very realistic simulator to achieve robust behavior. In this approach we are tuning a modified deterministic lookahead model (which is easy to solve) using a stochastic base model. This is fundamentally different than the approach that has become popular where you solve an approximation of a stochastic lookahead model.

It is not clear at this time how much energy from wind and solar can be absorbed into any particular system. This very much depends on access to fast-ramping reserves with very

short notification times. Our work has been done for PJM Interconnections, which covers the mid-Altantic states in the eastern U.S. This region has no access to hydroelectric reservoirs. Battery storage is in its infancy, and demand response markets generally need several hours of notification (although this is changing).

5.3 Risk Averse Approach with the SDDP Algorithm and a Brazilian Case Study (Alexander Shapiro, Georgia Tech)

5.3.1 Risk averse approach with the SDDP algorithm

In §4.3.3, the SDDP algorithm is applied to risk neutral problem (4.75), where the optimization (minimization) is performed *on average*. Of course, for a particular realization of the random data the costs $L_t(x_t, u_t, W_t)$ could be quite different from their expected (average) values. The goal of a risk averse approach is to avoid large values of the costs for some possible realizations of the data process at every stage of the considered time horizon.

One such approach will be to maintain constraints $L_t(x_t, u_t, W_t) \leq \gamma_t$, at every time period, for chosen upper levels γ_t and *all* possible realizations of the data process. However, trying to enforce these upper limits under any circumstances could be unrealistic and infeasible. One may try to relax these constraints by enforcing them with a high (close to one) probability. However, introducing such so-called chance constraints can still result in infeasibility and moreover is very difficult to handle numerically.

A way of dealing with intractability of chance constraints is to use a conservative approximation based on the Average Value-at-Risk measures

$$\mathsf{AV@R}_\alpha(Z) = \inf_{r \in \mathbb{R}} \left\{ r + \alpha^{-1}\mathbb{E}[Z - r]_+ \right\} \tag{5.10}$$

(cf., [42]). Minimum in the right hand side of (5.10) is attained at the quantile (also called Value-at-Risk in finance literature) $r^* = \mathsf{V@R}_\alpha(Z)$, where

$$\mathsf{V@R}_\alpha(Z) = F_Z^{-1}(1-\alpha) = \inf\{r : \mathsf{Pr}(Z > r) \leq \alpha\}$$

with $F_Z(z) = \mathsf{Pr}(Z \leq z)$ being the cdf of random variable Z. Moreover, if $F_Z(z)$ is continuous at $z = \mathsf{V@R}_\alpha(Z)$, then

$$\mathsf{AV@R}_\alpha(Z) = \mathbb{E}\left[Z | Z \geq \mathsf{V@R}_\alpha(Z)\right]. \tag{5.11}$$

In order to deal with the infeasibility problem we can move the constraints into the objective in a form of penalty. The corresponding dynamic programming equations are adjusted by replacing, at time period t, the expectation operator in the last equation of (4.48) with convex combination of the expectation and $\mathsf{AV@R}$ risk measure (cf., [57])

$$\rho_t(\cdot) = (1 - \lambda_t)\mathbb{E}(\cdot) + \lambda_t \mathsf{AV@R}_{\alpha_t}(\cdot). \tag{5.12}$$

Formulas (5.10) and (5.11) indicate that the $\mathsf{AV@R}_{\alpha_t}$ part of the risk measure ρ_t penalizes high quantiles of possible realizations of cost functions at time period t. That is, such risk averse approach tries to reach a compromise between optimizing (minimizing) on average and smoothing out possible high costs at every stage of the time process. The parameter $\lambda_t \in [0, 1]$ represents the weight given to the risk neutral and risk averse parts of the approach. Time consistency of this approach is ensures by the nested formulation of the problem.

The SDDP algorithm can be applied to the constructed risk averse problem with relatively small modifications. In the backward steps of the algorithm the corresponding cutting planes are computed by adjusting respective formulas for the subgradients of the risk measures of the value (cost-to-go) functions. The forward steps of the algorithm are performed basically in the same way as in the risk neutral case (see [66] for details).

In risk neutral implementations of the SDDP algorithm stopping criteria sometimes are based on estimation of the optimality gap of the constructed policy. Unfortunately for larger multistage problems this optimality gap does not become "small" in any reasonable computational time. That is, typically it is impossible to solve multistage stochastic programs with high accuracy accustomed in deterministic optimization. Realistic stopping criteria could be based on stabilization of the lower bound computed in the backward steps of the algorithm. The same approach can be used in the risk averse case.

5.3.2 Brazilian case study

We briefly discuss here case study [66] of the Brazilian interconnected power system. The Brazilian power system generation is hydro dominated (about 75% of the installed capacity) and characterized by large reservoirs presenting multi-year regulation capability, arranged in complex cascades over several river basins. The hydro plants use store water in the reservoirs to produce energy in the future, replacing fuel costs from the thermal units. Since the water inflows depend on rainfalls, the amount of future inflows is uncertain and cannot be predicted with a high accuracy.

The purpose of hydrothermal system operation planning is to define an operation strategy which, for each stage of the planning period, given the system state at the beginning of the stage, produces generation targets for each plant. The usual objective is to minimize the expected value of the total cost along the planning period, so as to meet requirements on the continuity of energy supply subject to feasibility constraints. The operation costs comprise fuel costs, purchases from neighboring systems and penalties for failure in load supply. This is referred to as the risk neutral approach: the total cost is optimized on average, and for a particular realization of the random data process the costs could be much higher than their average values. Risk averse approaches, on the other hand, aim at finding a compromise between minimizing the average cost and trying to control the upper limit of the costs for some possible realizations of the data set at every stage of the process.

The Brazilian hydro power operation planning problem is a multistage, large scale (more than 200 power plants, of which 141 are hydro plants), stochastic optimization problem. On a high level, planning is for 5 years on monthly basis together with 5 additional years to smooth out the end of horizon effect. This results in 120-stage stochastic programming problem. Four

energy equivalent reservoirs are considered, one in each one of the four interconnected main regions, SE, S, N and NE. The resulting policy obtained with the aggregate representation can be further refined, so as to provide decisions for each of the hydro and thermal power plants. In order to set the hydrothermal operating planning problem within this framework one can proceed as follows. Considering the aggregate representation of the hydroplants, the energy conservation equation for each equivalent energy reservoir n can be written as

$$SE_{t,n} = SE_{t-1,n} + CE_{t,n} - GH_{t,n} - SP_{t,n}. \tag{5.13}$$

That is, the stored energy (SE) at the end of each stage (start of the next stage) is equal to the initial stored energy plus controllable energy inflow (CE) minus total hydro generated energy (GH) and losses (SP) due to spillage, evaporation, etc.

At each stage, the net subsystem load L, given by the remaining load after discounting the uncontrolled energy inflow from the total load, has to be met by the total hydro, the sum of all thermal generation belonging to system n, given by the set NT_n, and the net interconnection energy flow (NF) to each subsystem. In other words, the energy balance equation for system n is

$$GH_{t,n} + \sum_{j \in NT_n} GT_{t,j} + NF_{t,n} = L_{t,n}. \tag{5.14}$$

Note that this equation is always feasible (i.e., the problem has relatively complete recourse) due to the inclusion of a dummy thermal plant with generation capacity equal to the demand and operation cost that reflects the social cost of not meeting the energy demand (deficit cost).

This leads to 120-stage linear stochastic programming problem with 8 state variables (4 additional state variables appear because of autoregressive modeling of the inflows process). The risk neutral and risk verse versions of the SDDP algorithm were applied to the constructed problem. Typical behavior of the upper and lower bounds of the SDDP algorithm are shown in Fig. 5.2.

Fig. 5.3 shows the individual stage costs for risk neutral and risk averse approaches for $\lambda = 0.15$ and $\alpha = 0.05$ and $\alpha = 0.1$. When we compare the risk averse approach and the risk neutral approach, we can see the significant reduction in the 99% quantile and the loss in the average policy value that occurs mostly in the first stages. Most of the reduction of the 90% and 95% quantiles happens in the last 15 stages. Furthermore, there is no significant difference between the individual stage costs for $\alpha = 0.05$ and $\alpha = 0.1$.

5.4 Transmission Grid Design (Alois Pichler, University of Vienna)

The electrical *power flow* in a (AC-) transmission network is described by a sequence of nonlinear equations. These equations can be solved given the demand and the supply in the network. While the demand in the companies and households is fixed, electricity can be

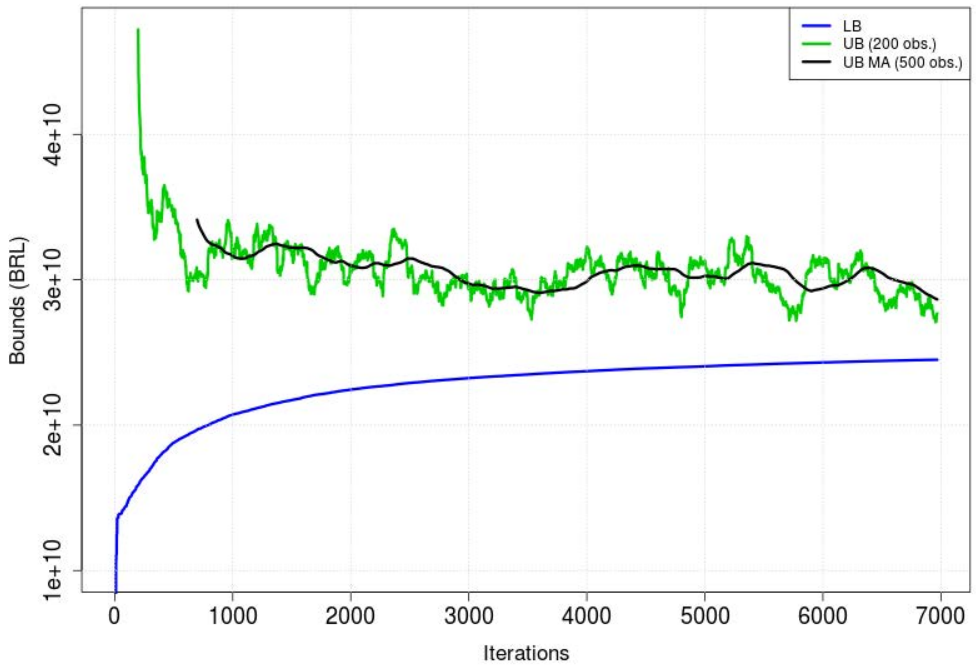


Figure 5.2: Upper and lower bounds, risk neutral case, 100 sample points per stage, 8 state variables, 120 stages, 1 cut per iteration

generated in different power plats in the network, as long as electricity can be transmitted to the demand nodes. To dispatch the production to different power plants can make sense, economicaly, as electric power can be produced at a different price in different power plants.

The *optimal power flow* problem addresses this question and intends to identify the amount of electricity, which should be generated in each particular plant.

Bienstock [15] points out that power flow problems can surprise optimization experts by their difficulty. For this reason they are often considered as benchmark problems, in particular for the choice of algorithms to numerically solve optimization problems.

From practical perspective, power flow problems are motivated by an increasing demand of electrical power and varying operational costs, as well as the need to manage system failures as blackouts (cf. Bienstock [15]).

Increasing importance has the need to incorporate renewable energy in the power flow network: many countries are in a transition period and currently redesigning their power network. Germany, for example, builds new transmission lines in order to transport electricity from the north, where electricity is generated in off-shore wind parks, to the south of the country, where demand is high. A new situation is given in Denmark (cf. Villumsen et al. [73]), as the country has decided that almost 50 % of its demand should be provided by wind power parks within a decade. These situations provide the opportunity to question and

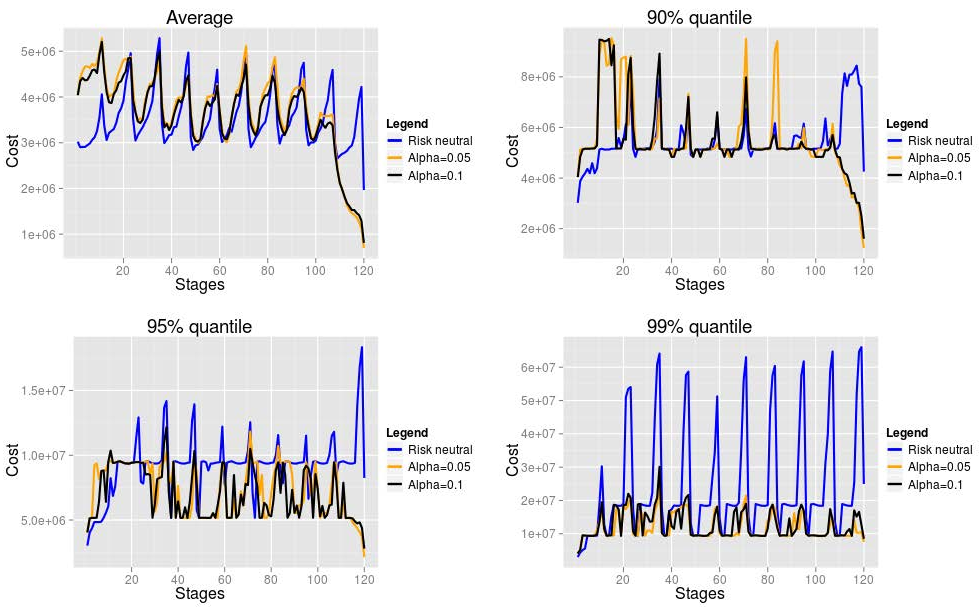


Figure 5.3: Individual stage costs for $\lambda = 0.15$ and $\alpha = 0.05$ and $\alpha = 0.1$.

rethink the existing power transmission network, and to improve its efficiency by establishing an optimal power network topology.

This exposes the optimal power flow problem in an economic, stochastic environment. The stochastic character is given by the fact that future demand and supply are random, they can only be assumed or estimated from today's perspective. However, the power grid has to be designed today, although its future utilization and the capacities necessary are not completely known today.

Villumsen and Philpott [74] consider the particular problem to decide on investments in electricity networks with transmission switching. They formulate the problem as a stochastic optimization problem by introducing scenarios. Every scenario ξ describes a reasonable pattern of demand and supply in the network, for which the power has to be generated at the costs $c(\cdot, \xi, \cdot)$.

These costs to generate the electric power are aggregated in a single objective function by assigning probability weights to every scenario ξ. In this way Villumsen and Philpott propose the two-stage stochastic optimization problem of the form

$$\min_{y \in Y} \mathbb{E}\left(\min_{z \in Z(y,\xi)} c\left(y, \xi, z\right) \right). \tag{5.15}$$

This model balances investment costs $y \in Y$ against potential reductions in operational cost $z \in Z$: the inner minimization models the decision of the system operator to generate and supply the power, which corresponds to the actual demand ξ in the network. It is the optimal power flow problem and the decision $z \in Z$, the *wait-and-see* decision, may differ for

different scenarios ξ. The expectation $\mathbb{E}$ summarizes the costs corresponding to ξ according the respective probability weight.

The *here-and-now* decision, $y \in Y$, identifies the optimal network design, which corresponds to an investment decision to be executed today. The investment decision consists in establishing the optimal grid topology by finding reasonable places to install expensive electrical equipment as FACTS devices (flexible AC transmission system) or automated switching opportunities.

It was observed that switching off existing lines may increase the overall efficiency and economic profitability of an electricity network (cf. Potluri and Hedman [50]). This fact (this paradox) is perhaps counter-intuitive and reflects the nonlinear properties of power flows, but it is clearly a starting point to incorporate switching possibilities in the network in order to adjust the power flow and to ensure constant, high profitability. This holds true even more as electricity demand, as well as its supply, are random: weather (as wind, rain or insolation), water-level of rivers and reservoirs, as well as outside temperature, time of day and time of the year are influencing factors amongst various others. Together with the expansion of renewable energy sources the respective volatility of electricity supply even increases. This situation creates a window of opportunity for power network design.

It was realized recently that the error caused by linearizing the inner problem in (5.15) exceeds the effect which is obtained by choosing different investment decisions $y \in Y$ (Fuller and Soroush in [27] point this out for switching in transmission systems). For this reason problem (5.15) has to be considered in its nonlinear formulation. A solution of the simplified, linear DC approximation is possibly a bad candidate for the genuine nonlinear problem (the AC formulation). Solving the linearized problem, even with high accuracy, thus does not allow justified conclusions to the real world problem.

For the investment problem (5.15) the inner optimization is nonlinear, whereas the outer minimization is combinatorial. The topology design problem thus is of *mixed integer, nonlinear, non-convex, combinatorial stochastic optimization* type. However, the evaluation of related functions is not expensive and analytic expression for derivatives are available. The optimal power flow problem in a stochastic environment thus nicely exposes the difficulties related to the nonlinear characteristics of (5.15) – a main reason why transmission switching was chosen to outline general solution techniques for optimization problems as (5.15).

5.5 Electricity Storage at End-User Level in a Smart Grid with Intermittent Local Supply (Stein W. Wallace, Zhan Pang, Pedro Crespo Del Granado, NHH, Bergen)

Our aim is to understand the role of electricity storage at end used level in a setting with local intermittent supply and a smart local grids. The analyses is set in a UK environment and we model and analyze both households and local communities. For households we use publicly

available consumption and wind data and our local community case study is the Lancaster University Campus. The campus has installed a CHP, three gas boilers, several micro-grid transformers and one wind-mill, but is considering one more. Our analysis concerns whether or not the university should also install a large battery to store electricity. The campus does not deliver electricity to the grid, so if local production (CHP plus wind) exceeds local demand, energy is lost. We use consumption data as well as wind data collected on site over several years.

We develop optimization models to analyze the value of energy storage in these two settings. The first paper([30]) analyzes the value of a battery in a private household. The second takes a more strategic view, and asks what is the value of a battery in a local community (in particular the campus) and how wrong the estimate can be if a deterministic strategic model is used. The third paper looks at the potential effect of such batteries on the behavior of the grid itself in case many end users install batteries, thereby changing the dynamics of the energy system. The second paper is submitted, the third is about to be submitted.

5.6 Benchmarking Electricity Markets Using Stochastic Optimization (Andy Philpott, Electric Power Optimization Centre, University of Auckland)

The last twenty years has witnessed a transformation in the development of market mechanisms for supplying electric power. In most western countries the centrally-planned systems developed during the twentieth century have been replaced by lightly regulated electricity markets. Market mechanisms aim to replace centrally-planned system optimizations by the individual optimizations of many agents. The goal of the market designer is to recover the system optimal solution but to do this by providing incentives rather than control. The incentives are intended to produce an optimal supply of power in the short run (from the cheapest plants) as well as providing sufficient profits to cover the long-run costs of generation, including the capital costs of any new plant, to now be provided by the private sector rather than governments. Ideally incentives will give a system optimal solution in both short and long run, but this is typically only possible in theory under some rather restrictive assumptions. For example, if the underlying short-run system optimization problem is not convex, then some convex approximation might be needed to enable it to be decomposed into agent problems.

5.6.1 Ex-ante market models

Most modeling questions about electricity market design are asked ex-ante. Market designers want to model a market mechanism before testing it in the real world when real money is at stake. Different mechanisms produce different overall efficiencies, but also different transfers of wealth between agents. Studying such mechanisms ex-ante uses game-theoretical models

of the market mechanisms to deduce the extent of market power that might arise in the implementation of such mechanisms.. In other words, by acting in their own self-interest, generators might seek to withhold generation to increase prices above competitive levels, leading to possible inefficiencies. Predictions of such outcomes can be made using equilibrium models of various types. It is common to seek a Nash equilibrium in which no generator wishes to unilaterally alter their generation level. There are many papers in the economics and applied mathematics communities that deal with such models in telecommunications, road traffic, supply chains etc. as well as electricity.

5.6.2 Market monitoring

After electricity markets are implemented, regulators seeks to study the performance of these by looking at market data. These data, even if not made public, should be available to the regulator. If observed prices are significantly above perfectly competitive benchmarks then a regulator has some evidence to support investigations of possible market manipulation. Social planning optimization models and the system marginal prices they yield (as Lagrange multipliers) are obvious candidates for competitive benchmarks. There are many econometric studies in the literature of this type (see e.g.[17], [80]).

Models of competitive electricity markets are relatively straightforward in a deterministic setting. Perfectly competitive markets consisting solely of thermal plant in a single period setting should dispatch these in order of increasing marginal cost. If these costs are known then it is easy to identify agents who misrepresent these.

Models with hydro

The marginal cost of supply is more difficult to identify for a hydro-dominated electricity market in which random inflows feed storage reservoirs. In this setting, a probabilistic model of the inflows yields an expected marginal water value for each reservoir that is the result of a stochastic optimization carried out by the agent. This value privides a bid price or opportunity value for releasing water from the reservoir.

If markets are complete and agents are risk neutral (maximizing expected profits) and perfectly competitive then one can show that competitive equilibrium corresponds to the optimal solution of a social planning problem that solves the stochastic control problem of meeting demand at least expected fuel and shortage cost. This problem can be solved at least approximately using dynamic programming methods such as stochastic dual dynamic programming (SDDP) [43]. The social planning policies can be simulated using historical inflow data and backtested against observed market outcomes. Some care is needed in interpeting results here as they are the outcome of testing a stochastic control policy using a single sample path. Even so, some observations can be made about the effiiency of the dispatch, price markups and wealth transfers (see e.g. [34],[49],[48] for papers adopting this approach).

Market incompleteness

There is an important caveat in using optimization models to benchmark electricity market performance: the socially optimal outcomes delivered in theory by a social planner do not always coincide with perfectly competitive equilibrium. If there are enough market instruments and agents are risk neutral, then an equilibrium under perfect competition will correspond to a socially optimal solution. Having enough market instruments is a *market completeness* assumption. A well-known example of market incompleteness was identified in Brazil when different agents owned generating stations at different points on the same river. As shown by Lino et al [38] this can lead to a loss of welfare in competitive equilibrium unless the market is completed with contracts to enable the agents to trade in the water they use for generating electricity.

Market incompleteness arising from risk

A more subtle form of incompleteness occurs from risk aversion. Here agents might believe the same probability distribution for uncertain events, but have different attitudes to risk. (We do not consider here the real possibility their probability distributions might actually differ, which leads to a different model.) Each agent's attitude to risk gives a higher weight to bad outcomes in a way that has been formalized by the theory of coherent risk measures (see [3],[55],[63]). Here the risk-adjusted cost of a random cost Z can be expressed as

$$\rho(Z) = \sup_{\mu \in \mathcal{D}} \mathbb{E}_\mu[Z],$$

in other words, the worst case expectation of the costs when taken with respect to some probability measure that lies in a convex set $\mathcal{D}$. When each agent maximizes their risk-adjusted profit, a competitive equilibrium is obtained that is often different from an optimal social plan. To see why, it is helpful to imagine what risk measure the social planner would use to account for the different preferences of all the different agents. When a hydro producer dislikes low reservoir inflows that yield low revenues, and a thermal producer dislikes high hydro reservoir inflows that yield low energy prices, then any coherent risk measure used by a social planner will fail to cover both these as worst cases.

Risk trading

To get closer to an optimal social plan, someone has to make some tradeoffs. This word is well chosen, since the tradeoffs can be shown to come from specific contracts traded between the players. In the example of the hydro reservoir the hydro agent arranges a contingent claim (a two-way option) with the thermal plant that pays the hydro player an amount from the thermal plant when reservoir inflows are low and a (possibly) different amount in reverse when reservoir inflows are high. The actual amounts paid settle based on the common probability of the dry event, and serve to reduce the risk of both parties. Such a contract was recently arranged in New Zealand between a thermal generator and a hydro generator (see [4]).

In the last few years a very elegant theory of market equilibrium and coherent risk measures has emerged through the work of several authors ([32],[54]). The key result established by Danny Ralph and Yves Smeers in [54] states that if each agent i has a coherent risk measure with risk set $\mathcal{D}_i$, and $\cap_i \mathcal{D}_i$ is polyhedral or has nonempty relative interior, and the market for trading risk is complete, then a risk averse social planning solution with risk set $\cap_i \mathcal{D}_i$ corresponds to a competitive equilibrium in which each agent optimally trades their risk. In this equilibrium all agents and the social planner view the random future through the same lens, placing higher weight on exactly the same scenarios. In other words the risk trading alters the players' payoffs so that they all agree on the worst case future scenarios. This result has been extended to a multistage setting (with finitely many scenarios) by [47].

This correspondence with a socially optimal solution enables regulators and market designers to go some way towards establishing good market mechanisms under perfect competition. The theorem indicates that there should be sufficient traded instruments to enable a social optimum to be attained. Numerical experiments with different mixtures of contracts indicate that some existing instruments are better than others, and can get most of the way to a socially optimal solution. These models are being used to understand how capital investment decisions are affected by risk in competitive markets [24], and how prices in markets with hydroelectricity rise as water shortages approach [47].

5.6.3 Challenges and open questions

Many mathematical challenges remain. Hydro-thermal scheduling problems with many reservoirs are high dimensional stochastic control problems. As mentioned above, approximate solutions to these can be obtained using variations of the stochastic dual dynamic programming (SDDP) algorithm [43], in both risk neutral and risk averse settings [62],[46]. This method constructs an outer approximation of the Bellman function of dynamic programming using Monte-Carlo sampling of the random variables and cutting planes. The solution to this multi-stage problem corresponds to a competitive equilibrium. This remains true in a scenario tree with risk averse agents trading in complete markets for risk at each node n with coherent risk measures satisfying $\cap_i \mathcal{D}_i(n) \neq \emptyset$. This enables regulators to postulate a dynamic risk measure that matches the intersection of individual risk measures and to benchmark actual market prices against those obtained with a competitive equilibrium and complete markets for risk.

The main challenge with such an approach comes from the computation. As mentioned above, computing approximate solutions to multi-stage social planning problems uses a dynamic programming approach that approximates the Bellman functions using cutting planes. This approach relies on the random processes having stagewise independent noise terms, a serious restriction. For most problems with no more than about ten reservoirs, the approximate Bellman functions converge to a policy that, when simulated, is provably close to optimal (with high probability). The subgradients of the Bellman function, however, show poorer convergence behaviour. This matters, as these marginal water values are estimates of equilibrium energy prices.

In practice, electricity markets are not complete. Then the optimization benchmark

will differ from a multi-stage competitive equilibrium, if this exists. Computing the latter remains a challenge that is actively being pursued by researchers in this area. In applications one needs to compute a competitive equilibrium at large scale, and so decomposition will be essential. Even if a competitive equilibrium with incomplete markets is not computable, the optimization benchmark can provide a bound on the welfare gain that could be obtained by enabling new financial instruments in the market being studied.

5.7 Probability Constraints in Dynamical Settings (Carpentier, Chancelier, De Lara, Leclère)

We consider the management over an one-year horizon of an hydroelectric dam subject to uncertain inflows, uncertain electricity prices and to a so-called "tourism constraint": the water storage level must be high enough during the tourist season with high enough probability. We formulate a chance-constrained stochastic optimal control problem, that is, we maximize the expected gain while guaranteeing a minimum storage level with a minimal prescribed probability level. Dualizing the chance constraint by a multiplier, we propose an iterative algorithm alternating Dynamic Programming and update of the multiplier value by a gradient step (see the dual gradient algorithm Algorithm in Remark 4.1). The numerical results reveal that the random gain is very dispersed around its expected value. In particular, low gain values have a relatively high probability to materialize. We conclude on the possibility to use the so-called stochastic viability approach, that focuses on jointly guaranteeing a minimum gain and a minimum storage level during the tourist season.

5.7.1 Chance constrained optimal management of a dam

Let $[\![0, T]\!] = \{0, 1, \dots, T\}$ represent the time span of the problem, and let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space. We consider the following real valued random variables:

- $\mathbf{X}_t$, the water *storage level* in the dam reservoir at the beginning of period $[t, t+1)$,
- $\mathbf{U}_t$, the dam *turbined outflow* during $[t, t+1)$,
- $\mathbf{A}_t$ and $\mathbf{C}_t$, the dam *inflow* and the *electricity price* during $[t, t+1)$.

We set $\mathbf{W}_t = (\mathbf{A}_t, \mathbf{C}_t)$ (*noise* at time t), and denote $\mathbf{W} = (\mathbf{W}_t)_{t \in [\![0, T-1]\!]}$ the noise process. We assume that $\mathbf{W}_0, \dots, \mathbf{W}_{T-1}$ are independent random variables.

The control strategy $\mathbf{U} = (\mathbf{U}_t)_{t \in [\![0, T-1]\!]}$ is required to be non-anticipative, that is,

$$\mathbf{U}_t \text{ is measurable w.r.t. } \sigma(\mathbf{W}_0, \dots, \mathbf{W}_t), \quad \forall t \in [\![0, T-1]\!], \tag{5.16}$$

where $\sigma(\mathbf{W}_0, \dots, \mathbf{W}_t)$ stands for the sigma-algebra generated by $(\mathbf{W}_0, \dots, \mathbf{W}_t)$. In addition, the control $\mathbf{U}_t$ cannot be greater than both the available water volume $\mathbf{X}_t + \mathbf{A}_t$ and the maximum turbined capacity $\overline{u}$:

$$0 \leq \mathbf{U}_t \leq \min\{\mathbf{X}_t + \mathbf{A}_t\,,\ \overline{u}\}, \quad \forall t \in [\![0, T-1]\!]. \tag{5.17}$$

We denote by $\mathfrak{U}$ the set of the control strategies that satisfy (5.16)—(5.17).

The dynamics of the storage level process $\mathbf{X} = (\mathbf{X}_t)_{t\in[\![0,T]\!]}$ reads

$$\mathbf{X}_{t+1} = F_t^{\mathbf{X}}\left(\mathbf{X}_t, \mathbf{U}_t, \mathbf{A}_t\right) = \min\left\{\mathbf{X}_t + \mathbf{A}_t - \mathbf{U}_t,\, \overline{x}\right\}, \quad \forall t \in [\![0, T-1]\!], \tag{5.18}$$

where $\mathbf{X}_0 = x_0$ is given and where $\overline{x}$ denotes the maximum water volume of the dam. Equation (5.18) describes a typical dam reservoir storage dynamics which takes into account the possible overflow of the dam.

Let the subset $\mathcal{T} \subset [\![1, T-1]\!]$ denote the tourist season period. The tourist chance constraint consists in ensuring a minimal reference storage level x_{ref} during the tourist season $\mathcal{T}$ at a probability level p_{ref}:

$$\mathbb{P}\left[\mathbf{X}_\tau \geq x_{\text{ref}} \;\; \forall \tau \in \mathcal{T}\right] \geq p_{\text{ref}}\,. \tag{5.19}$$

The gain obtained by producing electricity at time $t \in [\![1, T-1]\!]$ is modeled as

$$L_t(\mathbf{X}_t, \mathbf{U}_t, \mathbf{C}_t) = \mathbf{C}_t.\eta_t(\mathbf{X}_t, \mathbf{U}_t) - \epsilon\, {\mathbf{U}_t}^2\,, \tag{5.20}$$

where $\eta_t(\mathbf{X}_t, \mathbf{U}_t)$ is the electricity production at time t. The term $K(\mathbf{X}_T)$ represents the final value of water at time T which prevents the dam from being empty at the end of the horizon. The hydroelectric production gain of the dam is given by

$$\mathbf{G} = \sum_{t=0}^{T-1} L_t(\mathbf{X}_t, \mathbf{U}_t, \mathbf{C}_t) + K(\mathbf{X}_T)\,. \tag{5.21}$$

Summing up, we address the dam hydroelectric production management by the formulation

$$\max_{\mathbf{U}\in\mathfrak{U}} \quad \mathbb{E}\left[\sum_{t=0}^{T-1} L_t(\mathbf{X}_t, \mathbf{U}_t, \mathbf{C}_t) + K(\mathbf{X}_T)\right] \tag{5.22a}$$

$$\text{s.t.} \quad \mathbf{X}_{t+1} = F_t^{\mathbf{X}}\left(\mathbf{X}_t, \mathbf{U}_t, \mathbf{A}_t\right)\,, \tag{5.22b}$$

$$\mathbb{P}\left[\mathbf{X}_\tau \geq x_{\text{ref}} \;\; \forall \tau \in \mathcal{T}\right] \geq p_{\text{ref}}\,. \tag{5.22c}$$

The optimization problem (5.22) is a so-called chance constrained stochastic optimal control problem. Such problems raise theoretical difficulties: indeed, it is mathematically difficult to guarantee the connectedness, the convexity or the closedness of the feasible set induced by the chance constraint, although these properties play key roles in optimization. They also raise numerical difficulties, especially in the case of a joint chance constraint as in Problem (5.22).

5.7.2 Reformulation of the optimization problem and resolution

In order to deal with the joint probability constraint, we introduce the new binary valued random process $\boldsymbol{\pi} = (\boldsymbol{\pi}_t)_{t\in[\![1,T]\!]}$ driven by the dynamics

$$\boldsymbol{\pi}_{t+1} = F_t^{\boldsymbol{\pi}}\left(\mathbf{X}_t, \boldsymbol{\pi}_t, \mathbf{U}_t, \mathbf{A}_t\right) = \begin{cases} \mathbf{1}_{\{\mathbf{X}_{t+1}\geq x_{\text{ref}}\}}.\boldsymbol{\pi}_t & \text{if } \; t \in \mathcal{T} \\ \boldsymbol{\pi}_t & \text{else} \end{cases}, \quad \forall t \in [\![1, T-1]\!]\,, \tag{5.23}$$

with $\boldsymbol{\pi}_0 = 1$. Then, the chance constraint (5.19) can be formulated as

$$\mathbb{P}\left[\mathbf{X}_\tau \geq x_{\text{ref}} \ \ \forall \tau \in \mathcal{T}\right] = \mathbb{E}\left[\mathbf{1}_{\{\mathbf{X}_\tau \geq x_{\text{ref}} \ \ \forall \tau \in \mathcal{T}\}}\right] = \mathbb{E}\left[\boldsymbol{\pi}_T\right] \ ,$$

so that the optimization problem (5.22) reads

$$\max_{\mathbf{U} \in \mathfrak{U}} \quad \mathbb{E}\left[\sum_{t=0}^{T-1} L_t(\mathbf{X}_t, \mathbf{U}_t, \mathbf{C}_t) + K(\mathbf{X}_T)\right] \tag{5.24a}$$

$$\text{s.t.} \quad \mathbf{X}_{t+1} = F_t^{\mathbf{X}}\left(\mathbf{X}_t, \mathbf{U}_t, \mathbf{A}_t\right) \ , \tag{5.24b}$$

$$\boldsymbol{\pi}_{t+1} = F_t^{\boldsymbol{\pi}}\left(\mathbf{X}_t, \boldsymbol{\pi}_t, \mathbf{U}_t, \mathbf{A}_t\right) \ , \tag{5.24c}$$

$$\mathbb{E}\left[\boldsymbol{\pi}_T\right] \geq p_{\text{ref}} \ , \tag{5.24d}$$

with $\mathbf{X}_0 = x_0$ and $\boldsymbol{\pi}_0 = 1$.

We dualize the chance constraint (5.24d) by a multiplier λ which is a non negative scalar. The optimization problem (5.24) is equivalent to

$$\max_{\mathbf{U} \in \mathfrak{U}} \ \min_{\lambda \in \mathbb{R}_+} \ \mathbb{E}\left[\sum_{t=0}^{T-1} L_t(\mathbf{X}_t, \mathbf{U}_t, \mathbf{C}_t) + K(\mathbf{X}_T) + \lambda\left(\boldsymbol{\pi}_T - p_{\text{ref}}\right)\right] \tag{5.25}$$

$$\text{s.t.} \quad \text{(5.24b) and (5.24c)} \ ,$$

and an upper bound of the optimal value of Problem (5.25) is obtained by solving

$$\min_{\lambda \in \mathbb{R}_+} \ \max_{\mathbf{U} \in \mathfrak{U}} \quad \mathbb{E}\left[\sum_{t=0}^{T-1} L_t(\mathbf{X}_t, \mathbf{U}_t, \mathbf{C}_t) + K(\mathbf{X}_T) + \lambda\left(\boldsymbol{\pi}_T - p_{\text{ref}}\right)\right]$$

$$\text{s.t.} \quad \text{(5.24b) and (5.24c)} \ .$$

For a fixed value $\lambda^\sharp$ of the multiplier λ, we focus on the inner maximization problem

$$\max_{u \in \mathfrak{U}} \mathbb{E}\left[\sum_{t=0}^{T-1} L_t(\mathbf{X}_t, \mathbf{U}_t, \mathbf{C}_t) + K(\mathbf{X}_T) + \lambda^\sharp\left(\boldsymbol{\pi}_T - p_{\text{ref}}\right)\right] \tag{5.26}$$

$$\text{s.t.} \quad \text{(5.24b) and (5.24c)} \ .$$

If the strategy $\mathbf{U}^\sharp$ that solves the inner optimization problem (5.26) is such that the chance constraint (5.24d) is binding, that is, if

$$\mathbb{E}\left[\boldsymbol{\pi}_T^\sharp\right] = p_{\text{ref}} \ , \tag{5.27}$$

then Everett's theorem guarantees that the strategy $\mathbf{U}^\sharp$ is a solution to the optimization problem (5.24).

To find a multiplier value $\lambda^\star$ such that a solution $\mathbf{U}^\star$ to the inner optimization problem (5.26) satisfies the chance constraint (5.24d), we appeal to an iterative algorithm "à la Uzawa" (see the dual gradient algorithm Algorithm in Remark 4.1). This algorithm alternates two steps for each iteration k.

- *Inner maximization* (5.26): it is performed by solving the associated Dynamic Programming equation with state $(\mathbf{X}_t, \boldsymbol{\pi}_t)$ and leads, for the given multiplier value $\lambda^{(k)}$, to the computation of the strategy $\mathbf{U}^{(k+1)}$ and to the computation of the induced probability level $p^{(k+1)} = \mathbb{E}\left[\boldsymbol{\pi}_T^{(k+1)}\right]$.

- *Update of the multiplier*: the value $\lambda^{(k)}$ is updated by a gradient step method, namely,

$$\lambda^{(k+1)} = \max\left\{\lambda^{(k)} + \rho\left(p_{\text{ref}} - \mathbb{E}\left[\boldsymbol{\pi}_T^{(k+1)}\right]\right), 0\right\}, \quad \text{where} \quad \rho > 0\,. \tag{5.28}$$

It iterates until $\left(\mathbf{U}^{(k)}, \lambda^{(k)}\right)$ possibly converges to $(\mathbf{U}^\star, \lambda^\star)$. Note that the convergence proof of such an algorithm raises delicate issues because of the presence of the chance constraint.

5.7.3 Numerical experiments

We consider a dam management problem with the following features:

- tourist season: $\mathcal{T} = \{7, 8\}$, i.e. July and August months,
- tourist reference storage level: $x_{\text{ref}} = 50\ hm^3$,
- tourist reference probability level: $p_{\text{ref}} = 0.9$,
- maximum capacity of the dam reservoir: $\overline{x} = 80\ hm^3$,
- maximum water volume which can be turbined: $\overline{u} = 40\, hm^3$.

To solve the Dynamic Programming equation, the continuous spaces of states and controls are discretized by regular grids with $2\, hm^3$ steps. Regarding the implementation of the gradient step method, we set $\rho = 3,000$ in (5.28). The stopping criterion of the dual gradient algorithm is a maximal number of iterations of 800.

We see in Figure 5.4 that the dual gradient algorithm converges to a solution $(\mathbf{U}^\star, \lambda^\star)$ which saturates the chance constraint. Thus, $(\mathbf{U}^\star, \lambda^\star)$ solves the optimization problem (5.24) by virtue of Everett theorem.

We now focus of the optimal gain $\mathbf{G}^\star$ and on the optimal storage level process $\mathbf{X}^\star$. To depict these variables, we draw 10,000 realizations of the noise process $\mathbf{W}$ and we apply the optimal control strategy $\mathbf{U}^\star$ to each of these realizations. On Figure 5.5, we represent 100 trajectories (among the 10,000 obtained) of the optimal storage level process. We see that the tourist storage level x_{ref} is respected during the tourist season $\mathcal{T}$ for about 90% of the trajectories, which matches well the probability constraint. On Figure 5.6, we represent the empirical probability distribution of the optimal gain. We observe that the deviation of the random variable $\mathbf{G}^\star$ from its expected value is substantial: the standard deviation is about 40% of $\mathbb{E}[\mathbf{G}^\star]$, which might disappoint a dam manager who would expect a gain of the magnitude of its mean.

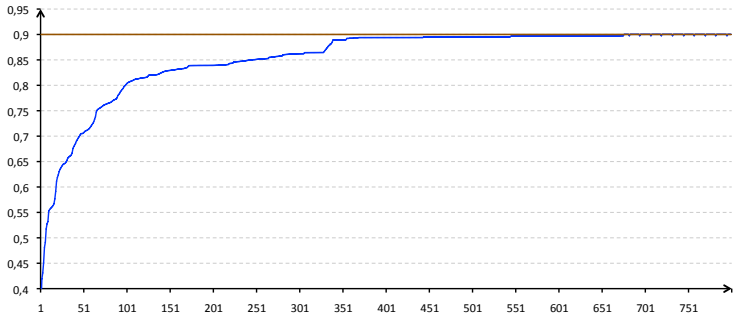


Figure 5.4: Evolution of the probability level $p^{(k)} = \mathbb{E}\left[\boldsymbol{\pi}_T^{(k)}\right]$ w.r.t. the iteration index k

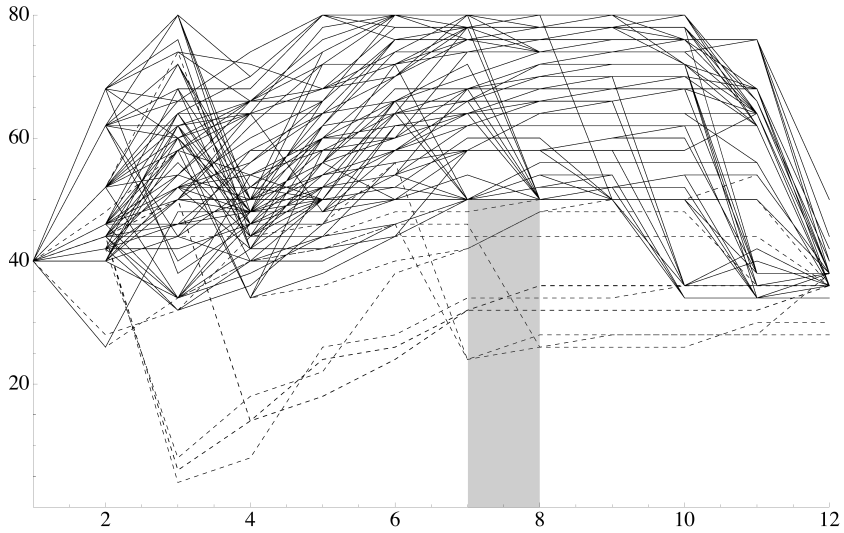


Figure 5.5: Realizations of the storage level process $\mathbf{X}^\star$; the dotted lines stand for the realizations that do not respect the tourist level x_{ref} during the tourist season $\mathcal{T}$

5.7.4 Conclusion and open problems

When moving from deterministic to stochastic control, modelers traditionally take the expected value of the original criterion (economists call this approach "risk-neutral"). They tackle constraints in various senses, such as robust, in probability one, or with a given probability level. Here, we chose to handle the dam management issue with the latter way. We considered a chance constrained stochastic optimal control problem, and we obtained satisfactory results with an algorithm that converged to an optimal solution. However, our numerical simulations revealed that the optimal random gain displayed a substantial dispersion.

Another way would be to use a stochastic viability approach that symmetrizes the economic and the tourist stakes, and jointly guarantees minimal thresholds. With this approach, a more complete picture of how to deal with the management of multi-purpose facilities under risk would be available.

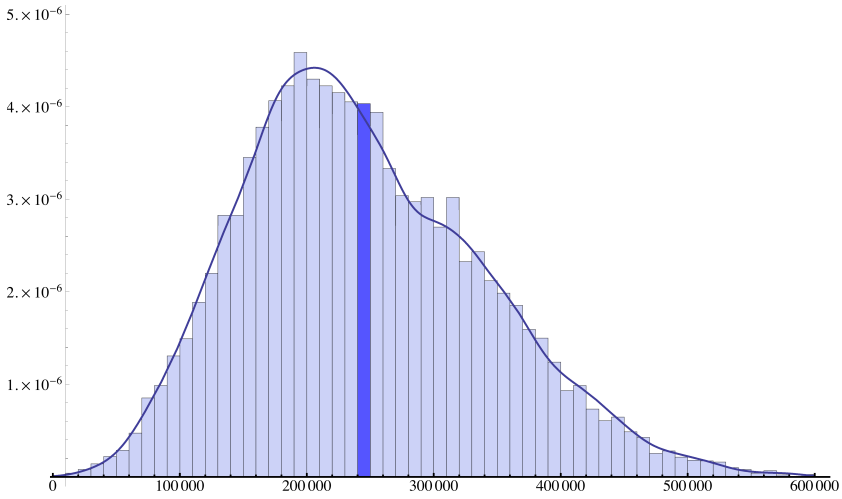


Figure 5.6: Empirical probability distribution of the gain $\mathbf{G}^\star$; the mean gain value is in the darkest box

Acknowledgements. The authors thank Électricité de France Research and Development for initiating this research through the CIFRE PhD funding of Jean-Christophe Alais and for supplying us with data.

5.8 Dual Approximate Dynamic Programming (DADP) (Carpentier, Chancelier, De Lara, Leclère)

We have presented in §4.1.2 a spatial decomposition method. We point out here some difficulties and an alternative approach known as *Dual Approximate Dynamic Programming*.

We consider N stochastic dynamic systems coupled by almost sure equality constraints. The global cost to be minimized is the expectation of a sum over the N systems of the sum over time of local costs. The problem considered will be detailed in §5.8.1. Our objective here is to obtain feedbacks (strategies).

The price decomposition scheme consists in dualizing the coupling constraints, fixing a multiplier, and obtaining N uncoupled subproblems. From the solution of each subproblem, we update the multiplier before iterating. However, we will show in §5.8.2 that this price decomposition scheme leads to subproblems which are too difficult to solve by Dynamic Programming (dimension of the state too important). Thus, we propose an approximation method based on the following ideas:

- relaxing the almost sure coupling equality constraints into conditional expectation constraint,
- using a price decomposition scheme to obtain subproblems,
- solving the subproblems through methods like Dynamic Programming.

The approximation idea behind the Dual Approximate Dynamic Programming (DADP) algorithm is presented in §5.8.3. A presentation of the scheme of DADP method is given in §5.8.4.

5.8.1 Presentation of the spatially coupled problem

We are interested in a production problem involving N units. Each unit i has an internal state $\boldsymbol{X}_t^i$ at time step t, and is affected by a random exogenous noise $\boldsymbol{W}^i$. The global exogenous noise $\{\boldsymbol{W}_t\}_0^{T-1}$ is assumed to be time-independent. Time dependence could be represented by extending the state, and incorporating information of the noise in it. On the other hand, for a given time t, the sequence $\{\boldsymbol{W}_t^i\}_{i=1}^N$ is not assumed to be independent (between units). Moreover, we assume a Hazard-Decision setting, that is, the control taken at time t is chosen once the uncertainty $\boldsymbol{W}_t$ is known.

For each unit $i \in [\![1, N]\!]$, and each time step $t \in [\![0, T-1]\!]$, we have to make a decision $\boldsymbol{U}_t^i \in \mathcal{U}_{t,i}^{ad}$ that must be measurable with respect to $\mathcal{F}_t$, where $\mathcal{F}_t$ is the σ-algebra generated by all past noises:

$$\mathcal{F}_t = \sigma\Big(\big(\boldsymbol{W}_s^i\big)_{1\leq i\leq N, 0\leq s\leq t} \Big) \ . \tag{5.29}$$

We denote $\mathfrak{F}$ the filtration $\{\mathcal{F}_t\}_0^{T-1}$.

We consider an almost sure coupling constraint represented as

$$\sum_{i=1}^N \theta_t^i\big(\boldsymbol{X}_t^i, \boldsymbol{U}_t^i, \boldsymbol{W}_t\big) = 0 \ , \ \ \mathbb{P}-a.s. \ . \tag{5.30}$$

For example, each θ_t^i can represent the production of unit i at time t, and a constraint on the global production at time t is represented through Equation (5.30).

Finally, the cost to be minimized is the expectation of a sum over time and over unit of all current local costs $L_t^i\big(\boldsymbol{X}_t^i, \boldsymbol{U}_t^i, \boldsymbol{W}_t\big)$.

The overall problem can be formulated as[2]

$$\min_{\boldsymbol{X},\boldsymbol{U}} \quad \mathbb{E}\Big[\sum_{i=1}^N \sum_{t=0}^T L_t^i\big(\boldsymbol{X}_t^i, \boldsymbol{U}_t^i, \boldsymbol{W}_t\big)\Big] \tag{5.31a}$$

$$\boldsymbol{X}_{t+1}^i = F_t^i(\boldsymbol{X}_t^i, \boldsymbol{U}_t^i, \boldsymbol{W}_t) \tag{5.31b}$$

$$\boldsymbol{U}_t^i \in \mathcal{U}_{t,i}^{ad} \tag{5.31c}$$

$$\boldsymbol{U}_t^i \preceq \mathcal{F}_t \tag{5.31d}$$

$$\sum_{i=1}^N \theta_t^i\big(\boldsymbol{X}_t^i, \boldsymbol{U}_t^i, \boldsymbol{W}_t\big) = 0 \tag{5.31e}$$

Note that, if it were not for constraint (5.31e), Problem (5.31) would lead to a sum of independent subproblems, that could be optimized independently.

[2]The final cost $L_T^i\big(\boldsymbol{X}_T, \boldsymbol{U}_T, \boldsymbol{W}_T\big)$ stands for a final cost of the form $K^i(\boldsymbol{X}_T)$ for the sake of notational simplicity.

5.8.2 First idea: price decomposition scheme

In §4.1.2, we presented how the dual gradient algorithm can be applied to a multistage problem. However, in §4.1.2 we did not specify how to solve the minimization problem for a given multiplier. Here, we use the dual gradient algorithm as the master problem in a price decomposition approach to Problem (5.31), and show its limits.

We dualize the coupling constraints (5.31e) to obtain

$$\max_{\boldsymbol{\lambda}\in \mathrm{L}^1} \quad \min_{\boldsymbol{X},\boldsymbol{U}} \quad \mathbb{E}\Big[\sum_{i=1}^{N}\sum_{t=0}^{T} L_t^i\big(\boldsymbol{X}_t^i,\boldsymbol{U}_t^i,\boldsymbol{W}_t\big) + \boldsymbol{\lambda}_t\cdot\theta_t^i\big(\boldsymbol{X}_t^i,\boldsymbol{U}_t^i,\boldsymbol{W}_t\big)\Big] \tag{5.32a}$$

$$\boldsymbol{X}_{t+1}^i = F_t^i(\boldsymbol{X}_t^i,\boldsymbol{U}_t^i,\boldsymbol{W}_t) \tag{5.32b}$$

$$\boldsymbol{X}_0^i = x_0^i \tag{5.32c}$$

$$\boldsymbol{U}_t^i \in \mathcal{U}_{t,i}^{ad} \tag{5.32d}$$

$$\boldsymbol{U}_t^i \preceq \mathcal{F}_t\ . \tag{5.32e}$$

Note that the multiplier $\boldsymbol{\lambda}$ is an adapted stochastic process $\boldsymbol{\lambda} = \big\{\boldsymbol{\lambda}_t\big\}_{t=0}^{T}$.

We can solve the maximization part of the dual problem using a gradient-like algorithm on $\boldsymbol{\lambda}$. Thus, for a fixed multiplier process $\boldsymbol{\lambda}^{(k)}$, we have to solve N independent problems of smaller size

$$\min_{\boldsymbol{X}^i,\boldsymbol{U}^i} \quad \mathbb{E}\Big[\sum_{t=0}^{T} L_t^i\big(\boldsymbol{X}_t^i,\boldsymbol{U}_t^i,\boldsymbol{W}_t\big) + \boldsymbol{\lambda}_t^{(k)}\cdot\theta_t^i\big(\boldsymbol{X}_t^i,\boldsymbol{U}_t^i,\boldsymbol{W}_t\big)\Big] \tag{5.33a}$$

$$\boldsymbol{X}_{t+1}^i = F_t^i(\boldsymbol{X}_t^i,\boldsymbol{U}_t^i,\boldsymbol{W}_t) \tag{5.33b}$$

$$\boldsymbol{U}_t^i \in \mathcal{U}_{t,i}^{ad} \tag{5.33c}$$

$$\boldsymbol{U}_t^i \preceq \mathcal{F}_t\ . \tag{5.33d}$$

Problem (5.31) is a multistage problem with a physical state $\boldsymbol{X}_t = \{\boldsymbol{X}_t^i\}_{i=1}^N$, affected by a time independent noise process $\{\boldsymbol{W}_t\}_0^T$. Hence, the state $\boldsymbol{X}_t$ is an information state in the sense of Dynamic Programming (see §4.1.4). Thus, Problem (5.31) can be solved through Dynamic Programming with a state of dimension $N \times dim(\mathbb{X}_t^i)$.

If it were not for the term $\boldsymbol{\lambda}_t^{(k)}\cdot\theta_t^i\big(\boldsymbol{X}_t^i,\boldsymbol{U}_t^i,\boldsymbol{W}_t\big)$ in the objective function of Problem (5.33), we would have a problem with a physical state $\boldsymbol{X}_t^i$ affected by the time independent noise process $\{\boldsymbol{W}_t\}_0^T$. Hence, the Dynamic Programming would be far faster: the dimension of the state is divided by N.

Unfortunately, with the term $\boldsymbol{\lambda}_t^{(k)}\cdot\theta_t^i\big(\boldsymbol{X}_t^i,\boldsymbol{U}_t^i,\boldsymbol{W}_t\big)$, Problem (5.33) is a problem with a physical state $\boldsymbol{X}_t^i$, and *two* random noises, namely $\boldsymbol{W}_t$ and $\boldsymbol{\lambda}_t^{(k)}$. The noise $\{\boldsymbol{W}_t\}_0^T$ process is time-independent, but the noise process $\{\boldsymbol{\lambda}_t^{(k)}\}_{t=0}^{T-1}$ has no reason whatever to be time independent. Hence, Dynamic Programming is not available with a reasonable information state.

If we could show that the multiplier process $\boldsymbol{\lambda}_t^{(k)}$ had a dynamic, say

$$\boldsymbol{\lambda}_t^{(k)} = h_t\big(\boldsymbol{\lambda}_{t-1}^{(k)},\cdots,\boldsymbol{\lambda}_{t-s}^{(k)},\boldsymbol{W}_t\big)\ , \tag{5.34}$$

then Problem (5.33) could be solved with the information state $\{\boldsymbol{X}_t^i, \boldsymbol{\lambda}_{t-1}^{(k)}, \cdots, \boldsymbol{\lambda}_{t-s}^{(k)}\}$. On a very specific example, it has been shown in [69] that the multiplier process has such a dynamic. In the following, we construct an approximation of Problem (5.31) such that its multiplier process is a function of a stochastic process $\boldsymbol{Y}$ with a dynamic. Our goal is to solve Problem (5.33) by Dynamic Programming with the extended information state $\big(\boldsymbol{X}_t^i, \boldsymbol{Y}_t\big)$.

5.8.3 Second idea: constraint relaxation

We have just seen that, if we apply a price decomposition scheme to Problem (5.31), the subproblems (5.33) cannot be solved numerically by the Dynamic Programming approach because of the curse of dimensionality. Thus, we approximate Problem (5.31) by relaxing the almost sure constraints, in order to obtain subproblems with a smaller dimension state, hence numerically solvable by Dynamic Programming.

For this purpose, we consider a stochastic process $\big\{\boldsymbol{Y}_t\big\}_{t=0}^{T-1}$ (uncontrolled), called an *information process*, that follows a dynamic

$$\boldsymbol{Y}_{t+1} = \widetilde{F}_t(\boldsymbol{Y}_t, \boldsymbol{W}_t) \ , \quad \forall t \in [\![0, T-1]\!] \ , \tag{5.35}$$

where $\widetilde{F}_t$ are known deterministic functions. The choice of the information process is arbitrary, but determines the quality of the method.

We replace, in Problem (5.31), constraint (5.31e) by its conditional expectation w.r.t the information process (see constraint (5.36e)):

$$\begin{aligned}
\min_{\boldsymbol{X}, \boldsymbol{U}} \quad & \mathbb{E}\Big[\sum_{i=1}^{N}\sum_{t=0}^{T} L_t^i(\boldsymbol{X}_t^i, \boldsymbol{U}_t^i, \boldsymbol{W}_t)\Big] && \text{(5.36a)}\\
& \boldsymbol{X}_{t+1}^i = F_t^i(\boldsymbol{X}_t^i, \boldsymbol{U}_t^i, \boldsymbol{W}_t) && \text{(5.36b)}\\
& \boldsymbol{U}_t^i \in \mathcal{U}_{t,i}^{ad} && \text{(5.36c)}\\
& \boldsymbol{U}_t^i \preceq \mathcal{F}_t && \text{(5.36d)}\\
& \mathbb{E}\Big[\sum_{i=1}^{N} \theta_t^i(\boldsymbol{X}_t^i, \boldsymbol{U}_t^i, \boldsymbol{W}_t) \ \Big| \ \boldsymbol{Y}_t\Big] = 0 \ . && \text{(5.36e)}
\end{aligned}$$

For this problem, we observe that we can restrict ourselves to multiplier processes $\boldsymbol{\mu}$ such that, for all time $t \in [\![0, T]\!]$, $\boldsymbol{\mu}_t$ is measurable w.r.t $\boldsymbol{Y}_t$, and we obtain

$$\begin{aligned}
\max_{\boldsymbol{\mu}_t \preceq \boldsymbol{Y}_t} \quad \sum_{i=1}^{N} \quad \min_{\boldsymbol{X}^i, \boldsymbol{U}^i} \quad & \mathbb{E}\Big[\sum_{t=0}^{T} L_t^i\big(\boldsymbol{X}_t^i, \boldsymbol{U}_t^i, \boldsymbol{W}_t\big) + \boldsymbol{\mu}_t \cdot \theta_t^i\big(\boldsymbol{X}_t^i, \boldsymbol{U}_t^i, \boldsymbol{W}_t\big)\Big] && \text{(5.37a)}\\
& \boldsymbol{X}_{t+1}^i = F_t^i(\boldsymbol{X}_t^i, \boldsymbol{U}_t^i, \boldsymbol{W}_t) && \text{(5.37b)}\\
& \boldsymbol{X}_0^i = x_0^i && \text{(5.37c)}\\
& \boldsymbol{U}_t^i \in \mathcal{U}_{t,i}^{ad} && \text{(5.37d)}\\
& \boldsymbol{U}_t^i \preceq \mathcal{F}_t \ . && \text{(5.37e)}
\end{aligned}$$

Note that, for a given multiplier $\boldsymbol{\mu}_t^{(k)}$, we have to solve the N following separate inner minimization subproblems.

$$\min_{\boldsymbol{X}^i, \boldsymbol{U}^i} \quad \mathbb{E}\Big[\sum_{t=0}^{T} L_t^i\big(\boldsymbol{X}_t^i, \boldsymbol{U}_t^i, \boldsymbol{W}_t\big) + \boldsymbol{\mu}_t^{(k)} \cdot \theta_t^i\big(\boldsymbol{X}_t^i, \boldsymbol{U}_t^i, \boldsymbol{W}_t\big)\Big] \tag{5.38a}$$

$$\boldsymbol{X}_{t+1}^i = F_t^i(\boldsymbol{X}_t^i, \boldsymbol{U}_t^i, \boldsymbol{W}_t) \tag{5.38b}$$

$$\boldsymbol{X}_0^i = x_0^i \tag{5.38c}$$

$$\boldsymbol{U}_t^i \in \mathcal{U}_{t,i}^{ad} \tag{5.38d}$$

$$\boldsymbol{U}_t^i \preceq \mathcal{F}_t \ . \tag{5.38e}$$

Each inner minimization problem can be solved by Dynamic Programming with the extended state $\big(\boldsymbol{X}_t^i, \boldsymbol{Y}_t\big)$. Indeed, fix a multiplier $\boldsymbol{\mu}_t^{(k)}$ measurable w.r.t $\boldsymbol{Y}_t$. A result from Doob (extended in [23, Chapter 1, p.18]) allows us to represent $\boldsymbol{\mu}_t^{(k)}$ as a function $\mu_t^{(k)}$ of $\boldsymbol{Y}_t$: $\mu_t^{(k)}(\boldsymbol{Y}_t) = \boldsymbol{\mu}_t^{(k)}$. Recalling that the noises $\{\boldsymbol{W}_t\}_0^{T-1}$ are assumed to be time-independent, we can write the following Dynamic Programming equation for the inner minimization problem:

$$V_t^i\big(x_t^i, y_t\big) = \min_{\boldsymbol{U}_t^i \preceq \boldsymbol{W}_t} \quad \mathbb{E}\Big[L_t^i\big(x_t^i, \boldsymbol{U}_t^i, \boldsymbol{W}_t\big) + \mu_t(y_t) \cdot \theta_t^i\big(x_t^i, \boldsymbol{U}_t^i, \boldsymbol{W}_t\big) + V_{t+1}^i\big(\boldsymbol{X}_{t+1}^i, \boldsymbol{Y}_{t+1}\big)\Big] \tag{5.39a}$$

$$\boldsymbol{X}_{t+1}^i = F_t^i(x_t^i, \boldsymbol{U}_t^i, \boldsymbol{W}_t) \tag{5.39b}$$

$$\boldsymbol{Y}_{t+1}^i = \widetilde{f}_t^i\big(y_t, \boldsymbol{W}_t\big) \tag{5.39c}$$

$$\boldsymbol{U}_t^i \in \mathcal{U}_{t,i}^{ad} \ . \tag{5.39d}$$

Thus, we can solve the inner minimization problem for a given multiplier, by applying Dynamic Programming to the N separate problems.

5.8.4 General scheme

We now describe more precisely the DADP Algorithm 5.1, given an information process $\big\{\boldsymbol{Y}_t\big\}_{t=0}^{T-1}$ satisfying (5.35).

Iteration k of Algorithm 5.1 starts with a multiplier process $\boldsymbol{\mu}_t^{(k)}$. The N inner minimization problems (5.38) are solved, for example, by Dynamic Programming. From these resolutions, we obtain a *slack process* $\boldsymbol{\Delta}_t^{(k)}$ defined by

$$\boldsymbol{\Delta}_t^{(k)} = \sum_{i=1}^{N} \theta_t^i\big(\boldsymbol{X}_t^{i,(k)}, \boldsymbol{U}_t^{i,(k)}, \boldsymbol{W}_t\big) \ , \tag{5.40}$$

where $\big\{\boldsymbol{X}_t^{i,(k)}, \boldsymbol{U}_t^{i,(k)}\big\}_{t=0}^{T}$ is the solution process of Problem (5.38).

Data: Information process evolution functions $\widetilde{F}_t$ and starting point y_0 initial multipliers $\boldsymbol{\mu}_t^{(0)}$;
Result: optimal multipliers $\mu_t^\star$, admissible feedback ;
repeat
  forall the $i \in [\![1, N]\!]$ **do**
    Solve Problem (5.38) ;
  forall the $t \in [\![0, T-1]\!]$ **do**
    Estimate $\mathbb{E}\big[\boldsymbol{\Delta}_t^k \mid \boldsymbol{Y}_t\big]$;
    Update the multiplier : $\boldsymbol{\mu}_t^{(k)}$ (5.41);
until $\mathbb{E}\big[\boldsymbol{\Delta}_t^k \mid \boldsymbol{Y}_t\big] \simeq 0$;
Compute admissible feedbacks ;

Algorithm 5.1: General Scheme of DADP

Then, we update the multiplier process by a gradient like step

$$\forall t \in [\![0, T]\!] \ , \ \ \boldsymbol{\mu}_t^{(k+1)} = \boldsymbol{\mu}_t^{(k)} + \rho \mathbb{E}\big[\boldsymbol{\Delta}_t^{(k)} \mid \boldsymbol{Y}_t\big] \ , \tag{5.41}$$

for a given $\rho > 0$. As $\boldsymbol{\mu}_t^{(k)}$ is measurable w.r.t $\boldsymbol{Y}_t$, if $\boldsymbol{Y}_t$ takes a finite number of values, $\mu_t^{(k)}$ can be represented by a finite dimensional vector. The update (5.41) of the multiplier process $\{\boldsymbol{\mu}_t^{(k)}\}_{t \in [\![0,T]\!]}$ ends iteration k of the DADP algorithm.

Chapter 6

Conclusion

The world's energy landscape is changing fast. In *Optimization Methods for the Smart Grid*, we have outlined three key drivers that are remolding power systems: renewable energies penetration, expansion of markets and of new players, deployment of telecommunication technology and smart meters. More precisely, the report sketches a smart grid paradigm. characterized by i) a growing number of decentralized and intermittent means of production (like solar and wind energies, the possible irregularity of which can complicate production management), ii) the deployment of smart meters and the development of automation of electric storage and electricity consuming devices (these intelligent and communicating equipments being supposed to make the electric consumption more controllable, hence more flexible).

We have highlighted how these changes impact and challenge the mathematical discipline of optimization. We have chosen a progressive and pedagogical exposition, mixing point of views from the energy sector and from the academic community, with the goal to make collaborations easier, especially regarding stochastic optimization. Thus, we enjoy the viewpoint of energy sector companies. In their contributions, they have exposed the trends they observe in their sector, and the possible new métiers and techniques they envisage.

Then, we have developed different mathematical formulations of optimization problems, especially under uncertainty, as solar and wind energies are intermittent, highly variable, and spatially scattered. For practitioners coming from deterministic optimization, this can be a substantial change of focus. Indeed, there is no single way to handle uncertainty when starting from a deterministic optimization formulation. Multiple framings are possible, depending on the attitudes of the decision-maker with respect to risk, and on the amount of information he or she has in hands at the making of a decision. To make the move more edible, we have worked out toy examples before laying out more general approaches.

With different mathematical formulations come different optimization methods. We have proposed a general umbrella for exposing methods from the literature. We stress that each of them emphasizes a special angle to break the original problem into more manageable pieces: decomposition with respect to time, to uncertainty (scenarios) or to space (units). We think that this approach is relevant, not only for exposition purpose, but also for future research. Indeed, we suggest that developing decentralized optimization methods — such as

decomposition and coordination methods — for problems where all the information (online) is not necessarily in the hands of a single decision-maker is especially adapted to the emerging context of energies management.

Speaking of research, we have asked academic colleagues to present their works. This is an opportunity to see, on the one hand, case studies and applications of methods already outlined, and, on the other hand, new developments that we could not tackle. Indeed, academic research in stochastic optimization is quite lively, fueled by internal demands and by the blossoming of applications. It was impossible to cover the whole field.

We thank all the academic colleagues and companies partners that have contributed to *Optimization Methods for the Smart Grid*. We stress that they are not engaged by the report's content. We hope that the material of the report can make academics and companies closer, when the emerging issues that we spotted in energy optimization turn more pressing.

Bibliography

[1] Roger N. Anderson, Albert Boulanger, Warren B. Powell, and Warren Scott. Adaptative stochastic control for the smart grid. In *Proceedings of the IEEE*, volume 99, issue 6, pages 1098–1115, June 2011.

[2] M.F. Anjos and J.B. Lasserre. *Handbook on Semidefinite, Conic and Polynomial Optimization*. Springer, 2012.

[3] P. Artzner, F. Delbaen, J.-M. Eber, and D. Heath. Coherent measures of risk. *Mathematical Finance*, 9:203–228, 1999.

[4] K. Barker. Meridian signs new hedge contract with Genesis. *NZ Energy News, Tuesday 8 July*, 2014.

[5] R. Bellman. Dynamic programming and the smoothing problem. *Management Science*, 3(1):111–113, 1956.

[6] R. E. Bellman. *Dynamic Programming*. Princeton University Press, Princeton, N.J., 1957.

[7] A. Bemporad and M. Morari. Control of systems integrating logic, dynamics, and constraints. *Automatica*, 35(3):407–427, March 1999.

[8] A. Ben-Tal, L. El Ghaoui, and A. Nemirovski. *Robust optimization*. Princeton University Press, 2009.

[9] A. Ben-Tal and A. Nemirovski. *Lectures on Modern Convex Optimization – Analysis, Algorithms, and Engineering Applications*. MPS-SIAM Series on Optimization 2. SIAM, 2001.

[10] A. Ben-Tal and A. Nemirovski. Robust optimization–methodology and applications. *Mathematical Programming*, 92(3):453–480, 2002.

[11] P. Bernhard. Max-plus algebra and mathematical fear in dynamic optimization. *Set-Valued Analysis*, 8(1-2):71–84, 2000.

[12] D. Bertsekas. *Dynamic Programming and Optimal Control*. Athena Scientific, Belmont, Massachusets, second edition, 2000. Volumes 1 and 2.

[13] D. Bertsekas. *Approximate dynamic programming*. Citeseer, 2011.

[14] D. Bertsekas and J. Tsitsiklis. *Neuro-Dynamic Programming*. Athena Scientific, Belmont, Massachusets, 1996.

[15] Daniel Bienstock. Progress on solving power flow problems. In *Optima 93*. Mathematical Optimization Society, 2013.

[16] J.R. Birge and F. Louveaux. *Introduction to stochastic programming*. Springer, 2011.

[17] S. Borenstein, J.B. Bushnell, and F.A. Wolak. Measuring market inefficiencies in California's restructured wholesale electricity market. *American Economic Review*, pages 1376–1405, 2002.

[18] A. Charnes and W.W. Cooper. Chance-constrained programming. *Management Science*, 6:73–79, 1959.

[19] G. Cohen. Auxiliary problem principle and decomposition of optimization problems. *Journal of optimization Theory and Applications*, 32(3):277–305, 1980.

[20] A. Dallagi. *Méthodes particulaires en commande optimale stochastique*. PhD thesis, Université Panthéon-Sorbonne-Paris I, 2007.

[21] G. Dantzig. *Linear programming and extensions*. Princeton university press, 1965.

[22] M. De Lara and L. Doyen. *Sustainable Management of Natural Resources. Mathematical Models and Methods*. Springer-Verlag, Berlin, 2008.

[23] C. Dellacherie and P.A. Meyer. *Probabilités et potentiel*. Hermann, Paris, 1975.

[24] A. Ehrenmann and Y. Smeers. Generation capacity expansion in a risky environment: a stochastic equilibrium analysis. *Operations Research*, 59(6):1332–1346, 2011.

[25] I. Ekeland and R. Temam. *Convex analysis and variational problems*. Society for Industrial and Applied Mathematics, 1999.

[26] H. Föllmer and A. Schied. *Stochastic Finance. An Introduction in Discrete Time*. Walter de Gruyter, Berlin, 2002.

[27] J. David Fuller and Milad Soroush. Accuracies of optimal transmission switching heuristics based on DCOPF and ACOPF. *IEEE Transactions on Power Systems*, 2013.

[28] P. Girardeau, V. Leclère, and A. Philpott. On the convergence of decomposition methods for multi-stage stochastic convex programs. *Optimization online, April*, 2012.

[29] D. E. Goldberg. *Genetic algorithms in search, optimization, and machine learning*, volume 412. Addison-wesley Reading Menlo Park, 1989.

[30] P. Granado, S. Wallace, and Z. Pang. The value of electricity storage in domestic homes: a smart grid perspective. *Energy Systems*, 5(2):211–232, 2014.

[31] T. Hastie, R. Tibshirani, and J. Friedman. *The elements of statistical learning*. Springer, 2009.

[32] D. Heath and H. Ku. Pareto equilibria with coherent measures of risk. *Mathematical Finance*, 14(2):163–172, 2004.

[33] Holger Heitsch and Werner Römisch. Scenario tree modeling for multistage stochastic programs. *Mathematical Programming*, 118(2):371–406, May 2009.

[34] O. Kauppi and M. Liski. An empirical model of imperfect dynamic competition and application to hydroelectricity storage. Technical Report 232, HECER, 2008.

[35] Alan J. King and Stein W. Wallace. *Modeling with Stochastic Programming*. Springer Series in Operations Research and Financial Engineering. Springer New York, 2012.

[36] J. K. Kok, M. J. J. Scheepers, and I. G. Kamphuis. Intelligence in electricity networks for embedding renewables and distributed generation. *Intelligent Systems, Control and Automation: Science and Engineering, Intelligent Infrastructures*, 42:179–209, 2010.

[37] D. Kuhn, W. Wiesemann, and A. Georghiou. Primal and dual linear decision rules in stochastic and robust optimization. *Mathematical Programming*, 130(1), 2011.

[38] P. Lino, L.A.N. Barroso, M.V.F. Pereira, R. Kelman, and M.H.C. Fampa. Bid-based dispatch of hydrothermal systems in competitive markets. *Annals of Operations Research*, 120(1):81–97, 2003.

[39] Aditya Mahajan and Demosthenis Teneketzisy. Multi-armed bandits. In *Foundations and Applications of Sensor Management*, pages 121–151. Springer-Verlag, 2007.

[40] L. B. Miller and H. Wagner. Chance-constrained programming with joint constraints. *Operations Research*, 13:930–945, 1965.

[41] G. Nemhauser and L. Wolsey. *Integer and combinatorial optimization*, volume 18. Wiley New York, 1988.

[42] A. Nemirovski and A. Shapiro. Convex approximations of chance constrained programs. *SIAM Journal on Optimization*, 17:969–996, 2006.

[43] M. V. F. Pereira and L. M. V. G. Pinto. Multi-stage stochastic optimization applied to energy planning. *Math. Program.*, 52:359–375, October 1991.

[44] Georg Ch. Pflug and Werner Römisch. *Modeling, measuring and managing risk*. World Scientific, Singapore, 2007.

[45] A. Philpott and Z. Guan. On the convergence of stochastic dual dynamic programming and related methods. *Operations Research Letters*, 36(4):450–455, 2008.

[46] A.B. Philpott, V.L. de Matos, and E.C. Finardi. On solving multistage stochastic programs with coherent risk measures. *Operations Research*, 2013.

[47] A.B. Philpott, M.C. Ferris, and R.J-B. Wets. Equilibrium, uncertainty and risk in hydrothermal electricity systems. Technical report, Electric Power Optimization Centre, www.epoc.org.nz, 2014.

[48] A.B. Philpott and Z. Guan. Models for estimating the performance of electricity markets with hydro-electric reservoir storage. Technical report, Electric Power Optimization Centre, University of Auckland, 2013.

[49] A.B. Philpott, Z. Guan, J. Khazaei, and G. Zakeri. Production inefficiency of electricity markets with hydro generation. *Utilities Policy*, 18(4):174–185, 2010.

[50] Tejaswi Potluri and Kory W. Hedman. Impacts of topology control on the ACOPF. In *Power and Energy Society General Meeting, 2012 IEEE*, pages 1–7, 2012.

[51] W. Powell. *Approximate Dynamic Programming: Solving the curses of dimensionality*, volume 703. John Wiley & Sons, 2007.

[52] Warren B Powell. What you should know about approximate dynamic programming. *Naval Research Logistics (NRL)*, 56(3):239–249, 2009.

[53] M. L. Puterman. *Markov Decision Processes*. Wiley, New York, 1994.

[54] D. Ralph and Y. Smeers. Pricing risk under risk measures: an introduction to stochastic-endogenous equilibria. *http://ssrn.com/abstract=1903897*, 2011.

[55] R.T. Rockafellar and S. Uryasev. Optimization of conditional value at risk. *The Journal of Risk*, 2:21–41, 2000.

[56] R.T. Rockafellar and R. J-B. Wets. Scenarios and policy aggregation in optimization under uncertainty. *Mathematics of operations research*, 16(1):119–147, 1991.

[57] A. Ruszczyński and A. Shapiro. Conditional risk mappings. *Mathematics of Operations Research*, 31:544–561, 2006.

[58] S. M. Ryan, R. J.-B. Wets, D. L. Woodruff, C. Silva-Monroy, and J.-P. Watson. Toward scalable, parallel progressive hedging for stochastic unit commitment. In *Power and Energy Society General Meeting (PES)*, pages 1–5. IEEE, 2013.

[59] M.P. Vecchi S. Kirkpatrick, C. D. Gelatt Jr. Optimization by simulated annealing. *Science*, 220(4598):671–680, 1983.

[60] L. J. Savage. *The Foundations of Statistics*. Dover, New York, second revised edition, 1972.

[61] F.C. Schweppe. Power: Power systems '2000': Hierarchical control strategies: Multilevel controls and home minis will enable utilities to buy and sell power at 'real time' rates determined by supply and demand. *Spectrum, IEEE*, 15(7):42–47, July 1978.

[62] A. Shapiro. Analysis of stochastic dual dynamic programming method. *European Journal of Operational Research*, 209:63–72, 2011.

[63] A. Shapiro, D. Dentcheva, and A. Ruszczynski. *Lectures on stochastic programming: modeling and theory*. The society for industrial and applied mathematics and the mathematical programming society, Philadelphia, USA, 2009.

[64] A. Shapiro and A. Nemirovski. On complexity of stochastic programming problems. In *Continuous optimization*. Springer, 2005.

[65] A. Shapiro, W. Tekaya, J. Paulo da Costa, and M. Pereira Soares. Final report for technical cooperation between Georgia Institute of Technology and ONS – Operador Nacional do Sistema Elétrico. Technical report, Georgia Institute of Technology, 2012.

[66] A. Shapiro, W. Tekaya, J. Paulo da Costa, and M. Pereira Soares. Risk neutral and risk averse stochastic dual dynamic programming method. *European Journal of Operational Research*, 224:375–391, 2013.

[67] Alexander Shapiro. On complexity of multistage stochastic programs. *Operations Research Letters*, 34(1):1–8, 2006.

[68] J Si, A G Barto, Warren B. Powell, and D Wunsch. Handbook of learning and approximate dynamic programming. *Wiley-IEEE Press*, 2004.

[69] C. Strugarek. *Approches variationnelles et autres contributions en optimisation stochastique*. PhD thesis, Ecole des Ponts ParisTech, 2006.

[70] S. Takriti, J. R. Birge, and E. Long. A stochastic model for the unit commitment problem. *Power Systems, IEEE Transactions on Power Systems*, 11(3):1497–1508, 1996.

[71] Aidan Tuohy and Peter Meibom. Unit commitment for systems with significant wind penetration. *Power Systems, IEEE Transactions on Power Systems*, 24(2):592–601, 2009.

[72] Hirofumi Uzawa. Iterative methods for concave programming. *Studies in linear and nonlinear programming*, 6, 1958.

[73] Jonas C. Villumsen, Geir Brønmo, and Andy B. Philpott. Line capacity expansion and transition switching in power systems with large-scale wind power. *IEEE Transactions on Power Systems*, 2012.

[74] Jonas C. Villumsen and Andy B. Philpott. Investment in electricity networks with transmission switching. *European Journal of Operational Research*, 222:377–385, 2012.

[75] Jean-Paul Watson and David L Woodruff. Progressive hedging innovations for a class of stochastic mixed-integer resource allocation problems. *Computational Management Science*, 8(4):355–370, 2011.

[76] L. Wehenkel, M. Pavella, E. Euxibie, and B. Heilbronn. Decision tree based transient stability method a case study. *Power Systems, IEEE Transactions on*, 9(1):459–469, Feb 1994.

[77] S B Thrun. The Role Of Exploration In Learning Control. In D A White, D A Sofge, and Eds. Handbook of Intelligent Control:. *Neural, Fuzzy and Adaptive Approaches. Van Nostrand Reinhold*, 1992.

[78] P. Whittle. *Optimization over Time: Dynamic Programming and Stochastic Control*, volume 1 and 2. John Wiley & Sons, New York, 1982.

[79] H. S. Witsenhausen. A standard form for sequential stochastic control. *Mathematical Systems Theory*, 7(1):5–11, 1973.

[80] F.A. Wolak. An assessment of the performance of the New Zealand wholesale electricity market. *Report for the New Zealand Commerce Commission*, 2009.

[81] H. Wolkowicz, R. Saigal, and L. Vandenberghe, editors. *Handbook of Semidefinite Programming – Theory, Algorithms, and Applications*. Kluwer Academic Publishers, 2000.